Vol. 64, No. 5 September, 1957

Psychological Review

THEODORE M. NEWCOMB, Editor
University of Michigan

CONTENTS

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STREETS, LANCASTER, PA.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

$8.00 volume $1.50 issue

Entered as second-class matter July 13, 1897, at the post office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Vol. 64, No. 5 September, 1957

THE PSYCHOLOGICAL REVIEW

THE PSYCHOLOGICAL SIGNIFICANCE OF THE CONCEPT OF "AROUSAL" OR "ACTIVATION"

ELIZABETH DUFFY

The Woman's College of the University of North Carolina

The concept of "arousal," "activation," or "energy mobilization," as developed by the writer over a period of many years (**7**, **9**, **10**, **11**, **13**), and employed by others in various contexts (**15**, **18**, **25**, **40**), has wide applicability in psychology.[1] A fuller discussion of the topic will be presented elsewhere. Pending its appearance, however, it may be of interest to point out some of the areas which this concept should serve to illuminate.

It has been argued in previous papers (**10**, **12**) that all variations in behavior may be described as variations in either the direction[2] of behavior or the intensity of behavior. Only one part of this argument is essential for the present purpose. Whatever may be the reaction to the attempt to reduce the descriptive categories of psychology to two basic types of concept, we can proceed without dispute provided only it is agreed that intensity is a characteristic of behavior which can be abstracted and studied separately. It is the intensity aspect of behavior which has been variously referred to as the degree of excitation, arousal, activation, or energy mobilization.

I have argued that such abstraction from the totality of behavior is a necessary procedure if the psychologist is to be enabled to manipulate variables in a way likely to provide solutions to some of his problems. Confusion of the direction of behavior with the intensity of behavior, resulting in their fortuitous combination in certain psychological concepts (**10**) and in the "trait" names used to describe personality (**12**), was suggested as a possible basis for some of the unrewarding findings in many psychological investigations. Since the intensity of response can vary independently of the direction of response, it was proposed that it should be measured independently and its correlates investigated.

Perhaps a parallel may be seen in the analysis of sensory function.[3] Before

[1] The terms "activation" and "arousal," as used here, do not refer specifically to the activation pattern in the EEG. On the contrary, they refer to variations in the arousal or excitation of the individual as a whole, as indicated roughly by any one of a number of physiological measures (*e.g.*, skin resistance, muscle tension, EEG, cardiovascular measures, and others). The degree of arousal appears to be best indicated by a combination of measures.

[2] "Direction" in behavior refers merely to the fact that the individual does "this" rather than "that," or responds positively to certain cues and negatively to others.

[3] For this suggestion of a parallel, I am indebted to Dr. R. B. Malmo, who, in the fall of 1955, was kind enough to read the major portion of my manuscript for a forthcoming book, and to discuss it with his staff.

progress could be made in the study of sensation and its physical correlates, it was necessary to separate the dimension of intensity from that of other sensory characteristics. In audition, for example, loudness was distinguished from pitch, and was related to a different type of variation in the physical stimulus. In vision, brightness was separated from hue, and each of these aspects of vision was related to the appropriate type of variation in the stimulus. Little progress in the understanding of sensation could have been made until suitable abstractions from the total sensory experience had been achieved, and these identifiable aspects of the totality had been investigated separately.

Measurement of the intensity of response (i.e., the degree of excitation, arousal, activation, or energy mobilization), it has been pointed out, may be achieved, at least in rough fashion, through various means (**9**, **10**, **13**, **15**). Among the physiological measures which may be employed are skin conductance, muscle tension, the electroencephalogram (EEG), pulse rate, respiration, and others. These measures show intercorrelations, although the correlation coefficients are not always high since there is patterning in the excitation of the individual, the nature of which appears to depend upon the specific stimulus situation and upon organic factors within the individual.[4] Nevertheless, there is evidence also of "generality" of the excitation. Hence a concept of arousal, or energy mobilization, appears to be justified.

It should be noted that the physiological measures which serve as indicants of arousal, and which correlate at least to some degree with each other, include measures of autonomic functions, of skeletal-muscle functioning, and of the functioning of the higher nerve centers. It is clear that it is the ***organism***, and not a single system, or a single aspect of response, which shows arousal or activation.

The historical roots of the concept of activation lie in Cannon's concept of "energy mobilization" during "emotion" (**3**). Unlike Cannon's concept, however, the present concept of activation or arousal is designed to describe the intensity aspect of *all* behavior (**10**, **12**). Referred to as the "degree of excitation," it was, in 1934, defined as "the extent to which the organism as a whole is activated or aroused" (**9**, p. 194). Both its definition and its proposed mode of measurement have in more recent publications followed the line suggested at that time (**10**, **13**). When, however, studies of the electroencephalogram provided data on the behavioral correlates of changes in the EEG, it was sugggested that this measure also provided an indication of the degree of arousal (**13**).

To those unfamiliar with the concept of activation, confusion frequently arises between the degree of internal arousal (referred to by the concept) and the vigor and extent of overt responses. While the degree of internal arousal usually correlates fairly closely with the intensity of overt response, a discrepancy between the two may be introduced by the intervention of inhibitory processes, a phenomenon which has not received the degree of attention to which it is entitled. An additional source of confusion is the tendency on the part of some to confuse activation or excitability with vitality. Actually, it is suggested that these two characteristics are more likely to be negatively related than to be positively related. The tendency to be frequently and intensely aroused

[4] The patterning of excitation is discussed more fully in the manuscript referred to in Footnote 3. It is believed that a more adequate concept of excitation, or activation, is thereby developed.

leads no doubt to fatigue and to a consequent reduction in vitality.

The chief point in regard to arousal, which I have repeatedly made (**10**, **11**, **12**, **13**), is that arousal occurs in a *continuum,* from a low point during deep sleep to a high point during extreme effort or great excitement, with no distinguishable break for such conditions as sleep or "emotion." Evidence supporting this contention has been presented specifically for skin conductance, muscle tension, and the EEG (**13**). Recently Lindsley has elaborated upon the conception as it applies to the EEG (**25**), although earlier, in his "activation theory of emotion" (**24**, pp. 504–509), he had been of the opinion that "emotion" and sleep were conditions which were correlated with certain changes in the EEG, while conditions intermediate between the two were held to be as yet unexplained.

The factors which produce variations in the degree of arousal are various. They include, apparently, drugs, hormones, variations in physical exertion, and variations in what is commonly referred to as the degree of motivation. It appears that differences in the degree of arousal in different individuals may have a genetic or an environmental basis, or both. This conclusion is suggested from animal studies and from the relatively few studies of human beings in which the problem has been considered.

One of the potential contributions to psychology of the concept of arousal is that of breaking down the distinction between "drives" or "motives" and "emotion" (**10**, **11**). The same kinds of physiological changes may be observed to occur in these variously designated conditions, and, depending upon the degree of arousal, to produce the same sorts of effect upon behavior. It has been contended that "emotion" is in no sense a unique condition, and that our investigations should not be directed toward the study of "emotion" as such (**9**).

In the study of "motivation," the concept of arousal is of distinct service. By means of the physiological measures which serve as indicants of arousal, we may secure a direct measure of the degree (intensity) of "motivation."[5] Any other measure must of necessity be less direct. When all factors affecting the level of arousal except the degree of incentive value or threat value are held constant, measurement of the degree of arousal affords a measure of the "motivating" value of a given situation. It also affords, incidentally, an objective measure of what is called the "stress" imposed by a situation.

Physiological measurements made in a wide variety of situations have shown the expected correspondence between the degree of arousal and the apparent degree of significance of the situation—i.e., its incentive value or its threat value (**13**). For example, men undergoing flight training were found to show more tension of the muscles during the solo stage of training than during other stages, and during the maneuvers of take-off and landing than during other maneuvers (**39**). Galvanic skin responses obtained during replies to questions about provocative social problems were found to be smaller if the replies were in harmony with group opinion than if they were not, and "Yes" responses were found in general to be associated with smaller galvanic reactions than "No" responses (**34**).

The concept of activation holds fur-

[5] The concept of "motivation," as currently employed, is a "compound" concept which incorporates a description of both the "drive level," or arousal aspect, of behavior and also the direction taken by behavior, i.e., the selectivity of response. These two aspects of behavior may vary independently, though both are characteristically affected by a certain stimulus-condition such as hunger.

ther significance for psychology by virtue of the fact that variations in the degree of activation are, on the average, accompanied by certain variations in overt response.[6] The degree of activation appears to affect the speed, the intensity, and the coordination of responses. In general, the optimal degree of activation appears to be a moderate one, the curve which expresses the relationship between activation and quality of performance taking the form of an inverted **U**. This conclusion, as it relates to muscular tension and performance, was suggested by me in 1932 (8, pp. 544–546), by Freeman in several papers published around that time (15), and later by Courts (4). That it holds also for other indicators of the degree of activation is suggested by Freeman's finding that skin resistance and reaction time, measured simultaneously on a single subject for 105 trials over a number of days, gave an inverted **U**-shaped curve when plotted on a graph (14). More recently the EEG has been found to show the same sort of relationship to reaction time (22).

The effect of any given degree of activation upon performance appears to vary, however, with a number of factors, including the nature of the task to be performed and certain characteristics of the individual—such as, perhaps, the ability to inhibit and coordinate responses under a high degree of excitation (8). Organismic interaction is the basic explanatory principle suggested to account for the particular effects upon performance of various degrees of activation. Such organismic interaction may also, it appears, have some effect upon sensory thresholds. Again the possibility presents itself that the relationship may take the form of an inverted **U**-shaped curve.

[6] These studies are reviewed in the manuscript referred to in Footnote 3.

When performance has been observed to vary under certain conditions, such as those of drowsiness, of fatigue, or of "emotion," it is suggested that the variation may be due, at least in part, to the effect of varying degrees of arousal. The disorganization of responses frequently reported during "overmotivation" or "emotion," for example, may be conceived of as resulting in part from too high a degree of arousal. Such a condition would be represented at one end of the **U**-shaped curve. A similar disorganization of responses, found sometimes during drowsiness or fatigue, would be represented at the other end of the curve showing the relationship between arousal and performance. In any case, it seems clear that prediction of overt response to a given set of stimulating conditions can be increased in accuracy when there is knowledge of the degree of internal arousal.

It appears also that, under similar stimulation, individuals differ in the degree of their arousal and in the speed with which they return to their former level of functioning. Moreover, there is evidence of consistency in this individual variation. Apparently the individual who responds with intensity in one situation will, on the average, respond with intensity in other situations also, as compared with other individuals. While the degree of arousal varies with the situation, the rank in arousal tends to be preserved. Different individuals appear to vary around different central tendencies —i.e., to differ in responsiveness. The easily aroused, or more responsive, individual has been found to show this responsiveness in many different forms, some of which will be described below.

For instance, subjects who showed a large number of galvanic skin responses when there was no observable stimulation also showed less adaptation of the galvanic skin response (GSR) to repeated stimulation (33).

Similarly, the frequency of the alpha rhythm in normal adults has been reported to show a significant relationship to ratings on the behavioral continuum called "primary-secondary function" (32). Individuals in whom the alpha rhythm was more rapid tended to show more "primary functioning," or to be "quick, impulsive, variable, and highly stimulable." Those with relatively low frequencies of the alpha rhythm tended to show more "secondary functioning," or to be "slow, cautious, steady, with an even mood and psychic tempo. . . ." Mundy-Castle hypothetically ascribed these behavioral differences to differences in excitability within the central nervous system, the "primary functioning" individuals showing the greater excitability. A difference in neural excitability was also suggested as the explanation of his finding that there was a significant difference in the EEG activity evoked by rhythmic photic stimulation between subjects with a mean alpha frequency above 10.3 cycles per second and those with a mean alpha frequency below that rate.[7] He offered the same explanation of the greater incidence of "following"[8] in the beta range by those subjects showing little alpha rhythm, even when the eyes were closed, as compared with those subjects showing persistent alpha rhythms (32).

Gastaut and his collaborators have also reported individual differences in cortical excitability (17). While their major purpose was not the investigation of individual differences, they made the incidental observation that calm individuals had a slow, high-voltage alpha rhythm (8–10 c./s.), with little "driving" of occipital rhythms by photic stimulation. Neurons showed a long recuperation time, synchrony of response was said to be noticeable, and recruitment poor. "Nervous" individuals, on the other hand, were said to have a high-frequency, low-voltage alpha rhythm (10–13 c./s.), which at times was not perceptible. They were described as having a short neuronal recuperation time, little synchrony of response, good recruitment, and considerable driving by photic stimulation. In other words, "calm" as compared with "nervous" individuals showed less cortical excitability.

Differences in the EEG's of different individuals under similar stimulating conditions appear to be correlated also with differences in another form of responsiveness—i.e., differences in the threshold of deep reflexes. It has been reported that normal subjects with deep reflexes which are difficult to elicit showed a high percentage of alpha activity and little or no fast activity, while those with deep reflexes which were hyperactive had little alpha activity and a high percentage of fast activity (21). However, while groups at the two extremes of reflex responsiveness differed significantly in the percentage of alpha activity, there was wide variation in the extent of such activity within any one of the groups formed on the basis of reflex status. Amplitude of rhythm was observed to be greatest in EEG records showing pronounced alpha activity.

Proneness to develop anxiety under stress, which may perhaps be regarded as a form of hyperresponsiveness, has been found, in both normal subjects and psychiatric patients, to be associated with a significantly smaller percentage

[7] It is believed, he says, that "electrical rhythms in the brain can be initiated or augmented by a process similar to resonance; in other words, if an area of the brain is subjected to rhythmic impulses corresponding to its own latent or actual frequency, it may itself oscillate for as long as stimulation is maintained" (33, p. 319). It is thought that the area may also be activated by stimulation harmonically related to its own.

[8] "Following" refers to electrical responses in the cortex occurring at the stimulus frequency.

of resting brain-wave activity in the alpha region when this activity is determined by automatic frequency analysis (35). The anxiety-prone groups showed more fast activity (16–24 c./s.), or more slow activity (3–7 c./s.), below the alpha range. The significance of the slow activity is not as clear as that of the fast activity. Fast activity may be presumed to be indicative of a high level of excitation. It has been observed, for example, at the beginning of EEG recording in normal subjects who are unusually apprehensive about the procedure, and it has been found to disappear with reassurance and the attainment of relaxation (24). It appears at least possible that the slow activity may be due to fatigue from previous states of intense arousal.

In an investigation employing prison farm inmates, schizophrenics, and control subjects, to whom a group of psychological tests were given, it was reported that EEG activity above 16–20 c./s. appeared in significant amounts only in the records of those who, as rated by the psychological tests, showed anxiety to a marked degree (20). Slow activity was said not to be very prevalent, but when it did occur, to be found most often among the patients.

These and other studies suggest that anxiety-proneness may be conceived of as a form of overarousal or hyperresponsiveness. The EEG's of the anxiety-prone seem very similar in most instances to the EEG's of other subjects whose exceptional responsiveness to the environment is indicated by active reflexes, or by ratings on "primary function."

Degree of tension of the skeletal muscles is another indicator of responsiveness, or ease and extent of arousal, in which differences between individuals have been found. In almost every investigation in which tension of the skeletal musculature has been measured, wide differences between individuals in the degree of tension have been noted.[9] In the same stimulus situation, one individual would respond with a relatively low degree of tension, another with a moderate degree, and a third with a high degree of tension. Moreover, when observed in a *different* stimulus situation, the subjects, while varying in their absolute level of tension, would tend to preserve their ranks with respect to tension of the muscles. It was thus shown that different individuals vary around different central tendencies, so that one individual might be characterized as being in general tense, and another as being in general relaxed.

In early studies of muscular tension, the writer found, in two separate investigations, that nursery school children showed marked individual differences in grip pressure while engaged in various tasks, and that there was a significant correlation between the grip pressure on one occasion and that on another, and during one task and during another (6, 7). Grip pressure scores were found to be independent of the strength of grip as indicated by dynamometer scores, but to be related to ratings on excitability and on adjustment to the nursery school, the tense children being rated as more excitable and, on the average, less well adjusted.

Arnold also found that individuals tended to preserve their rank in the group with respect to pressure from the hand during repetition of the same task and during the performance of different tasks (2).

A study of airplane pilots in training revealed that some showed excessive muscle tension (pressure on stick and on rudder pedal) in both take-offs and

[9] Differences in muscle tension will, for the purposes of this discussion, refer to differences in pressure exerted by some group of muscles or to differences in electric potentials from muscles.

landings, while others showed little tension on either maneuver (39). No individuals were found who in general tended to be tense during take-offs alone or during landings alone.

Further evidence that individuals who are more highly activated than others in one stimulus situation, as indicated by tension of the skeletal muscles, are more responsive to a wide variety of stimuli, is presented in studies by Lundervold (26). "Tense" subjects, as compared with "relaxed" subjects, were found to show more activity in the muscles when external conditions were changed, as by an increase in noise, the lowering of the room temperature, or the introduction of certain stimuli which caused irritation or anger. In these persons, there was not only more activity in the single muscle, but also electrical activity in more muscles, including muscles which did not participate directly in the movement. At the end of thirty minutes of noise, fifty per cent of the tense subjects, as compared with none of the relaxed subjects, showed more action potentials than they had shown before the noise began.

A similar relationship between muscular tension and another form of responsiveness was earlier shown by Freeman and Katzoff, who found a significant correlation between grip-pressure scores and scores on the Cason Common Annoyance Test (16). Subjects with higher pressure scores tended to be more frequently or intensely annoyed—i.e., to show greater responsiveness of the sort referred to as "irritability."

It appears that, on the whole, skeletal-muscle tension in one part of the body tends to be positively related to that in other parts of the body, though the relationship between the tension in any two areas may not be very close. Parts of the body more remote from each other, or more widely differentiated in function, yield tension measures which are less closely related than those which are closer together or functionally more similar. When tension measures taken from different parts of the body, recorded during different tasks, or made at widely separated intervals of time, nevertheless show a significant positive correlation with each other, it must, however, be concluded that there is at least some degree of "generality" in skeletal-muscle tension. Moreover, from measuring the responsiveness of the skeletal-muscle system, we may apparently predict to some extent the response of highly integrated systems of reaction described as "personality traits." Indeed, in a study in which no direct measure of muscular tension was employed, but in which ratings on muscular tension and measures of sixteen physiological variables were intercorrelated and submitted to factor analysis, a factor defined as muscular tension showed correlation with certain personality characteristics (36).

Since conditions of high activation may perhaps increase the likelihood of disorganization of motor responses, it is not surprising that measures of tremor and other forms of motor disorganization have been found to be related to the severity of conflicts (31) and to neuroticism (1, 19, 23, 28, 29, 30). Measures of irregularity in pressure appear to be among the measures which discriminate best between a normal and a psychiatric population, a finding which might be expected if, as suggested by the writer (8) and by Luria (27), irregular pressure tracings are indicative of poor coordination or lack of control of responses.

Other indicants of arousal have also been shown to be related to more complex forms of response. For example, it has been said that a reasonably accurate prediction of a person's respiratory rate at a given time during a flight could be made on the basis of knowl-

edge of his "normal" respiratory rate and the name of the maneuver to be performed (**39**).

Similarly, when an "autonomic factor" was obtained from twenty physiological measures related to the functioning of the autonomic nervous system, it was found that individuals differed greatly in scores on this factor, but that the correlation coefficient between early and later factor scores did not drop below .64 over a two-year period (**38**). Children at one extreme of the autonomic-factor scores were reported to differ significantly from those at the other extreme in certain personality traits (**37**).

Individuals differ, not only in the degree of excitation produced by stimulation, but also in the speed with which the processes affected return to their prior level of functioning. Moreover, differences in recovery time cannot be accounted for solely by differences in the degree of arousal, for they are found when recovery is measured *in relation to the degree of arousal.* Darrow and Heath, who first made use of this measure, computed a "recovery-reaction quotient" by dividing the extent of recovery in skin resistance by the extent of decrease in resistance which had occurred as a result of stimulation (**5**). The recovery-reaction quotient was reported to be related to many different measures of " 'neurotic' and emotionally unstable tendencies." The investigators concluded that it was one of their best indicators of the absence of neurotic trend, but that the coefficients of correlation were not high enough to justify the use of the measure for prediction in individual cases. It would appear that the speed of recovery from arousal is an extremely significant aspect of response, and one which deserves further investigation.

Individuals who are exceptionally responsive to the environment may show their responsivity in behavior which, from a directional point of view, may be described in diverse ways. A tendency toward a high degree of arousal does not determine which aspects of the environment an individual will approach or will have a tendency to approach (i.e., have a favorable attitude toward); nor does it determine which aspects of the environment he will withdraw from or have a tendency to withdraw from (i.e., have an unfavorable attitude toward). On the contrary, the orientation of the individual in his environment is determined largely by other factors. These are, of course, the factors, both genetic and environmental, which have given to various aspects of his environment the nature of their significance, or their "cue-function." There are, nevertheless, differences in the way in which approach or withdrawal occurs which may conceivably be derived from differences in the level of activation. Among these appear to be differences in such aspects of behavior as alertness, impulsiveness, irritability, distractibility, and the degree of organization of responses. Moreover, greater responsiveness may, it is suggested, facilitate the development of aggression or withdrawal, enthusiasm, or anxiety. The more responsive individual in a certain kind of environment is no doubt more susceptible to the effects of that environment. Presumably he may become, depending upon circumstances, more anxiety-prone, more conscientious, more sympathetic, more devoted, or more irascible than a less responsive person would become under similar circumstances. We should therefore expect to find some association between a high degree of activation and easily aroused or intense responses of various kinds (e.g., anxieties, resentments, enthusiasms, or attachments). From knowledge of the individual's tendencies with respect to activation we should not, however, be able to predict the direction which his behavior

would take. A more dependable association might be expected between individual differences in excitability and differences in the "dynamic" characteristics of behavior such as those mentioned above.

The effect of a high degree of arousal upon overt behavior varies, no doubt, with variations in the degree of inhibitory ability (**9**), or, as Luria has described it, with variations in the strength of the "functional barrier" between excitation and response (**27**).[10] Depending upon this factor, a high degree of activation may, I suggest, lead to impulsive, disorganized behavior or to sensitive, alert, vigorous, and coordinated responses to the environment. Evidence in support of these statements is at present so meager, however, as to leave them in the category of speculations. It is to be hoped that further investigation will provide the basis for a more confident statement of the relationship between "personality" characteristics and individual differences in the level of activation.

SUMMARY

The concept of arousal or activation appears to be a significant one for the ordering of psychological data. Differences in activation, as shown in a wide variety of physiological measures, appear to be associated with many other differences in response.

In different stimulus-situations, the same individual differs in the degree of arousal. Measurement of the physiological indicants of arousal affords, when other factors are constant, a direct measure of the "motivating" or "emotional" value of the situation to the individual. The concept serves to break down the distinction between the arousal aspect of "drives" or "motives" and that of "emotion," and to suggest instead a continuum in the degree of activation of the individual.

Differences in activation in the same individual are, it is suggested, accompanied by differences in the quality of performance; the relationship may be graphically represented by an inverted U-shaped curve. Further data are needed, however, to establish the validity of this hypothesis.

In the same stimulus situation there are differences between individuals in the degree of arousal. These differences tend to persist, and thus to characterize the individual. Moreover, the easily aroused, or responsive, person shows this responsiveness in many forms. It has been observed in the ease with which deep reflexes are elicited, and in the extent, frequency, and duration of reactions to stimulation, both of the skeletal musculature and of various functions controlled by the autonomic nervous system. It has been shown also in differences in cortical potentials, which are presumably indicative of differences in the excitability of higher nerve centers. These various forms of responsiveness show, in general, positive intercorrelations, though the coefficients of correlation are apparently not high enough for a measure of any one mode of responsiveness to serve as an adequate measure of the general responsivity of the individual. They appear, however, to give justification to the conception of a responsive or an unresponsive *individual,* not merely responsive or unresponsive skeletal musculature, skin resistance, or cortical potentials.

Differences in arousal are shown also in responses of greater inclusiveness and

[10] Luria reports that children show weakness of the functional barrier between excitation and motor response, as indicated by poor performance on a test requiring that a key be pressed down as slowly as possible (**28**). The writer noted that, during a discrimination performance, younger nursery school children, with irregular grip-pressure tracings, had a higher proportion of their errors in the category of "impulsive" errors, or errors of overreaction (**8**).

of higher integration—i.e., in responses frequently classified as personality traits. Combining with one or another directional aspect of behavior, a persistent high degree of arousal may, it appears, be observed in many complex characteristics, such as anxiety-proneness or aggressiveness.

Facts such as those presented above suggest that the concept of activation may prove useful in many different areas of psychology.

REFERENCES

1. Albino, R. C. The stable and labile personality types of Luria in clinically normal individuals. *Brit. J. Psychol.*, 1948, **39**, 54–60.
2. Arnold, M. B. A study of tension in relation to breakdown. *J. gen. Psychol.*, 1942, **26**, 315–346.
3. Cannon, W. B. *Bodily changes in pain, hunger, fear and rage.* New York: Appleton, 1915, 1929.
4. Courts, F. A. Relations between muscular tension and performance. *Psychol. Bull.*, 1942, **39**, 347–367.
5. Darrow, C. W., & Heath, L. L. Reaction tendencies related to personality. In K. S. Lashley (Ed.), *Studies in the dynamics of behavior.* Chicago: Univer. of Chicago Press, 1932. Pp. 59–261.
6. Duffy, E. Tensions and emotional factors in reaction. *Genet. Psychol. Monogr.*, 1930, **7**, 1–79.
7. Duffy, E. The measurement of muscular tension as a technique for the study of emotional tendencies. *Amer. J. Psychol.*, 1932, **44**, 146–162.
8. Duffy, E. The relationship between muscular tension and quality of performance. *Amer. J. Psychol.*, 1932, **44**, 535–546.
9. Duffy, E. Emotion: an example of the need for reorientation in psychology. *Psychol. Rev.*, 1934, **41**, 184–198.
10. Duffy, E. The conceptual categories of psychology: a suggestion for revision. *Psychol. Rev.*, 1941, **48**, 177–203.
11. Duffy, E. An explanation of "emotional" phenomena without the use of the concept "emotion." *J. gen. Psychol.*, 1941, **25**, 283–293.
12. Duffy, E. A systematic framework for the description of personality. *J. abnorm. soc. Psychol.*, 1949, **44**, 175–190.
13. Duffy, E. The concept of energy mobilization. *Psychol. Rev.*, 1951, **58**, 30–40.
14. Freeman, G. L. The relationship between performance level and bodily activity level. *J. exp. Psychol.*, 1940, **26**, 602–608.
15. Freeman, G. L. *The energetics of human behavior.* Ithaca: Cornell Univer. Press, 1948.
16. Freeman, G. L., & Katzoff, E. T. Muscular tension and irritability. *Amer. J. Psychol.*, 1932, **44**, 789–792.
17. Gastaut, H. et Y., Roger, A., Corriol, J., & Naquet, R. Étude électrographique du cycle d'excitabilité cortical. *EEG clin. Neurophysiol.*, 1951, **3**, 401–428.
18. Hebb, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.
19. Jost, H. Some physiological changes during frustration. *Child Develpm.*, 1941, **12**, 9–15.
20. Kennard, M. A., Rabinovitch, M. S., & Fister, W. P. The use of frequency analysis in the interpretation of the EEG's of patients with psychological disorders. *EEG clin. Neurophysiol.*, 1955, **7**, 29–38.
21. Kennard, M. A., & Willner, M. D. Correlation between electroencephalograms and deep reflexes in normal adults. *Dis. nerv. System*, 1943, **6**, 337–347.
22. Lansing, R. W., Schwartz, E., & Lindsley, D. B. Reaction time and EEG activation. *Amer. Psychologist*, 1956, **11**, 433.
23. Lee, M. A. M. The relation of the knee jerk and standing steadiness to nervous instability. *J. abnorm. soc. Psychol.*, 1931, **26**, 212–228.
24. Lindsley, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 473–516.
25. Lindsley, D. B. Psychological phenomena and the electroencephalogram. *EEG clin. Neurophysiol.*, 1952, **4**, 443–456.
26. Lundervold, A. An electromyographic investigation of tense and relaxed subjects. *J. nerv. ment. Dis.*, 1952, **115**, 512–525.
27. Luria, A. R. *The nature of human conflict* (Transl. and ed. by W. H. Gantt). New York: Liveright, 1932.
28. Malmo, R. B., Shagass, C., Bélanger, D. J., & Smith, A. A. Motor control in psychiatric patients under experimen-

tal stress. *J. abnorm. soc. Psychol.,* 1951, **46**, 539–547.

29. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.,* 1951, **12**, 45–66.
30. MALMO, R. B., & SMITH, A. A. Forehead tension and motor irregularities in psychoneurotic patients under stress. *J. Pers.,* 1955, **23**, 391–406.
31. MORGAN, M. I., & OJEMANN, R. H. A study of the Luria method. *J. appl. Psychol.,* 1942, **26**, 168–179.
32. MUNDY-CASTLE, A. C. Electrical responses of the brain in relation to behavior. *Brit. J. Psychol.,* 1953, **44**, 318–329.
33. MUNDY-CASTLE, A. C., & MCKIEVER, B. L. The psychophysiological significance of the galvanic skin response. *J. exp. Psychol.,* 1953, **46**, 15–24.
34. MURRAY, H. A. *Explorations in personality.* New York: Oxford Univer. Press, 1938.
35. ULETT, G. A., GLESER, G., WINOKUR, G., & LAWLER, A. The EEG and reaction to photic stimulation as an index of anxiety-proneness. *EEG clin. Neurophysiol.,* 1953, **5**, 23–32.
36. WENGER, M. A. An attempt to appraise individual differences in level of muscular tension. *J. exp. Psychol.,* 1943, **32**, 213–225.
37. WENGER, M. A. Preliminary study of the significance of measures of autonomic balance. *Psychosom. Med.,* 1947, **9**, 301–309.
38. WENGER, M. A., & ELLINGTON, M. The measurement of autonomic balance in children: method and normative data. *Psychosom. Med.,* 1943, **5**, 241–253.
39. WILLIAMS, A. C., JR., MACMILLAN, J. W., & JENKINS, J. G. *Preliminary experimental investigations of "tension" as a determinant of performance in flight training.* Civil Aeronautics Admin., Div. of Res., Rep. No. 54, Washington, D. C. January, 1946.
40. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology.* (Rev. ed.) New York: Holt, 1954.

(Received October 22, 1956)

Psychological Review
Vol. 64, No. 5, 1957

ANXIETY AND BEHAVIORAL AROUSAL[1]

ROBERT B. MALMO[2]

Allan Memorial Institute of Psychiatry, McGill University

During the past two decades there has been a growing interest in objective physiological studies of psychiatric patients. In this work, one of the most prominent psychological concepts has been that of anxiety. Although there is general agreement that the areas denoted by the term "anxiety" are important ones for study, there is nonetheless considerable disagreement concerning what the term means. In large measure, this semantic difficulty is part of a larger problem facing psychology today, and that is to find a way out of the confusion surrounding the concepts of motivation and emotion. Duffy has cogently argued that these concepts are second-order ones which reduce to primary factors of intensity and direction, and that along the intensity dimension, at least, the distinction between motivation and emotion is unnecessary (**9, 10, 11**).[3]

This is not to say that the directional aspect is not important or to deny that, in terms of direction, meaningful distinctions may be made between motivation and emotion, and indeed between different emotions. However, for present purposes it is essential to focus on the question of what these phenomena have in common rather than to consider their differences; in this paper, therefore, we shall be primarily concerned with the intensity dimension.

The main purpose of the present paper is to consider recent experimental data in an attempt to find a way out of the present confusion. I shall begin with a summary of two lines of investigation in our laboratory, dealing first with our discovery that certain physiological measures may serve as indicants of intensity or "behavioral arousal." These experiments were performed with nonpatient subjects. Second, in summarizing our investigations of pathological anxiety in psychiatric patients, I shall attempt to use the concept of behavioral arousal in an integrative way. Third, I shall draw on data from recent neurophysiological investigations to indicate possible mechanisms involved in the pathology and etiology of anxiety. Finally, on the basis of these theoretical considerations, I suggest problems requiring further experimental study.

Physiological Indicants of Behavioral Intensity

In 1951 we (**31**) reported finding a gradient phenomenon from electromyographic (EMG) recording during mirror tracing. Since that time the phenomenon has been observed under various conditions in our laboratory. Figure 1 presents mirror-drawing data from a study by Bartoshuk (**1**). Note that the

[1] This paper reviews work which was supported by the Medical Research and Development Division, Office of the Surgeon General, Department of the U. S. Army, under Contract Number DA 49-007-MD-626, by Defence Research Board Grant Number 9425-04 (Canada), and by Grant Number A.P. 29 from the National Research Council of Canada.

[2] The author is indebted to Drs. A. K. Bartoshuk, D. Bindra, F. R. Brush, D. E. Cameron, D. O. Hebb, and R. G. Stennett for criticizing earlier drafts of this paper.

[3] I do not wish to imply that this has been Duffy's only theoretical contribution. Her writings contain prior reference to a dimension of behavioral intensity (conceived as a continuum of "arousal," or "activation"); and she has previously cited evidence to support the argument that physiological measures may serve as the chief means of quantifying such a dimension or continuum.

chin lead (which taps the speech muscles) also shows a gradient—that is, progressively rising muscle potentials from the beginning to the end of the task. Bélanger (3) found similar gradients from the arm in a size-discrimination task. Wallerstein (42) found gradients in the frontalis muscle in a task about as completely devoid of motor components as one could possibly design. The subject, reclining on a comfortable bed, listened to verbal material (short detective story or essay) presented to him by a tape recorder. In Wallerstein's experiment, the gradients extended over ten minutes and their steepness was related to the subject's reported degree of interest in listening (2, p. *228* f.).

Bartoshuk (2) was the first to show that the fastest and most accurate subjects (i.e., superior performers on mirror tracing) produced the steepest muscle-potential gradients. Such a relationship of EMG gradients to motivation has been confirmed by three subsequent studies, employing tracking tasks. Surwillo (39) demonstrated that raising incentive had the effect of increasing the steepness of EMG gradients in a visual tracking experiment. Figure *2* presents confirmatory data from a more recent experiment by Stennett (37) who employed auditory tracking under four conditions, with increasing degrees of incentive. Note that the muscle potentials were recorded from the nonactive, left arm. His "exertion" condition merely involved the subject's holding the tracking knob over at a fixed point in order to control for sheer physical work. Under the "calibration" condition the subject believed that he was just assisting with calibration of the apparatus, and that his tracking scores were not being recorded. The "optimal" condition was designed to motivate the subject sufficiently to elicit his most efficient performance, whereas the "incentive" condition was designed to "overmotivate" the subject by offering large bonuses for high-level performance and threatening with strong electric shock if performance did not reach this high level. The differences shown in the figure were statistically significant. In brief, Stennett's findings indicated that the most efficient tracking performance was associated with intermediate physiological levels (i.e., intermediate steepness of EMG gradients and intermediate levels of palmar skin conductance). With lower levels of physiological functioning (less steep gradients, lower levels of palmar skin conductance), performance on tracking was inferior. However, going now to the other extreme, performance on tracking associated with extremely high EMG gradients and extremely high palmar skin conductance was also inferior to tracking performance associated with moderately high levels of physiological functioning.

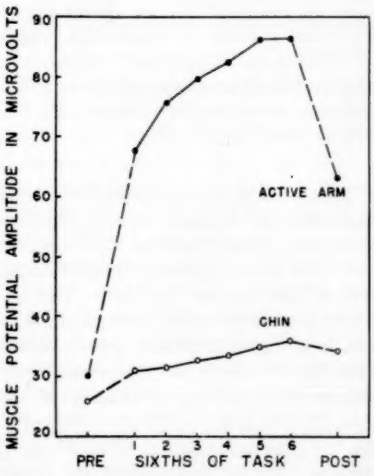

FIG. 1. Graphs showing mean EMG gradients in Bartoshuk's experiment (1). Note that gradient was also obtained from chin lead which records from muscles of speech. $N = 17$.

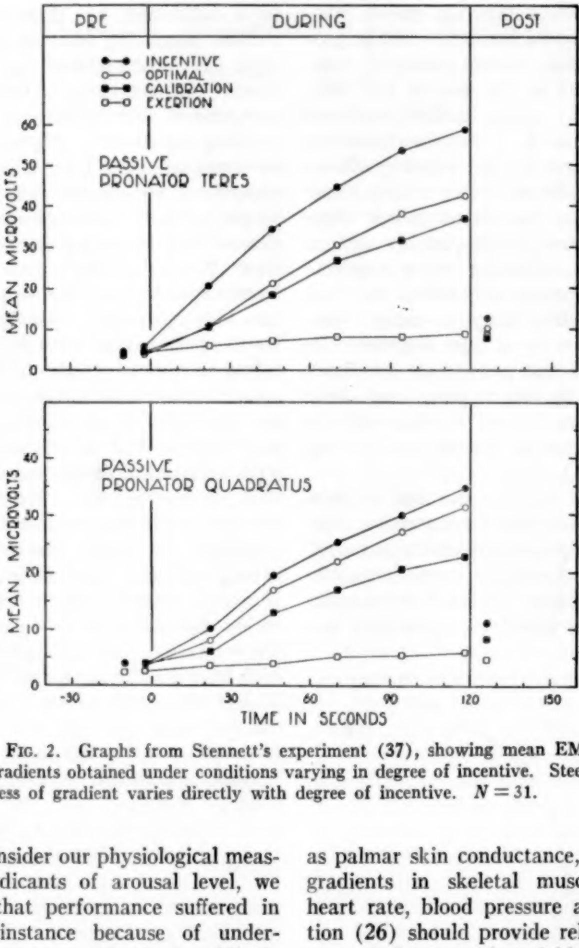

FIG. 2. Graphs from Stennett's experiment (37), showing mean EMG gradients obtained under conditions varying in degree of incentive. Steepness of gradient varies directly with degree of incentive. $N = 31$.

If we consider our physiological measures as indicants of arousal level, we may say that performance suffered in the first instance because of underarousal (or poor motivation), while in the second instance it suffered from overarousal (or emotional interference). In short, as Stennett has previously stated (37), we believe that the concept of arousal leads us in the direction of working out (empirically) a continuum of behavioral intensity which promises to have the very desirable feature of integrating the concepts of motivation and emotion. From available data it appears that physiological measures, such as palmar skin conductance, EEG [4] and gradients in skeletal muscle tension, heart rate, blood pressure and respiration (26) should provide reliable measures of the arousal variable. The objective nature of the physiological measures is a highly desirable feature which frees the worker from dependence upon merely manipulating situations in the

[4] Stennett (38) has found that the relationship of alpha amplitude to arousal level is nonlinear. On the lower end of the arousal continuum the relationship is positive, such that raising arousal leads to increasing alpha amplitude; but past the middle range of arousal the relationship becomes inverse. This latter function is the better known one.

hope that he is producing intended changes in the arousal level of his subjects. Moreover, the physiological indicants have the further advantage that they may be applied to work with animals as well as with human subjects, and may thus serve usefully to bridge the gap, in the field of motivation, between work on human and on infrahuman subjects.

A word should be said concerning the different physiological measures which have served as indicants of behavioral arousal. Although gradient steepness has proved a very useful measure, level of palmar skin conductance seems equally promising. As a matter of fact, even with EMG, the correlation between average EMG level and gradient steepness is usually so high that it is meaningless to ask which is a better indicant. We still have much to learn concerning the application of physiological techniques to our problems. It may be that, as Lacey's work suggests (**23**), for most accurate assessment of arousal, special consideration should be given to individual differences in relative reactivity of different physiological systems.

Following the usage of Freeman (**15**) and Hebb (**18**), the term "arousal" is used to refer to the intensive dimension. I am aware that the term "arousal" is used by some EEG workers to denote flattening of an EEG tracing (e.g., **8**, p. 132). When I use the term, I use it in a much broader sense, as a dimension of behavior, and I am not using this term to refer to the EEG phenomenon called "arousal" or "activation." It is for this reason that I specify *behavioral* arousal in the title of this paper. As investigative work proceeds, it may become heuristic to make a definite distinction between physiological arousal and behavioral intensity. Granting this possibility, I believe that for present purposes it may be preferable to accept a rather broad operational definition of the intensity dimension, in which level of physiological activity, arousal, and intensive level are employed as roughly synonymous terms.

In short, the physiological measures appear to be useful tools in establishing and precisely quantifying a dimension of behavioral intensity. Indeed, I regard such objective measures as nearly indispensable to the achievement of a really satisfactory operational definition of behavioral intensity. In the absence of such objective measures, it is difficult to see how circularity can be avoided. Considerable work is required, of course, in working out the intensity dimension, and while present results are indeed encouraging, many further data are required. It may be helpful just here to relate the arousal continuum to the intensity dimension which Boring described (**4**). While Boring's main concern was with sensation, I believe that it is appropriate to consider that operations of measurement comparable in precision to those of psychophysics may be possible in the field of action.

EXPERIMENTAL STUDIES WITH PSYCHIATRIC PATIENTS

Having elucidated the concept of arousal with these reference experiments, we are now in a favorable position to take a fresh look at the data comparing patients and nonpatients with respect to level of physiological reaction under controlled stimulating conditions. At the outset, we may say that the chief impression which one gets in going over all of these data is that, under "stress," psychoneurotic patients appeared to show a higher level of physiological reaction than controls, and that level of reaction seemed particularly high in patients suffering mainly from pathological anxiety. By pathological anxiety, I mean a state of such severity that work efficiency is seriously affected over long periods of time, and a state which is

characterized by one or more of the following complaints: persistent feelings of "tension" or "strain," "irritability," "unremitting worry," "restlessness," "inability to concentrate," "feelings of panic in everyday-life situations." I should like to make it very clear that I do not employ the term "anxiety" to refer to transient affective states. When I use the term I am talking about a pathological condition which—as far as we can determine—develops slowly, over months or years, and from which recovery (when it occurs) is also slow and gradual. The experiments which we shall consider in this section employed patients suffering from "pathological anxiety," as we have just defined it. For the sake of convenience, these subjects will be called "anxiety patients."

In a study with pain as standard stimulus (**27**), the following physiological measures showed significantly greater reaction in anxiety patients than in other psychiatric patients: finger movement (and number of voluntary pressures to indicate pain), neck-muscle activity, deviation in amplitude and rate of respiration throughout the test, respiratory irregularities occurring at time of stimulation, and heart-rate variability. In a different study (**29**) with a perceptual test and a Luria-type recording from the left hand, finger movement was significantly more irregular in anxiety patients than in other psychiatric patients.

To repeat an earlier statement, these findings indicate that under standard conditions of stimulation psychoneurotics are more reactive than controls, and that patients with anxiety predominating in the symptom picture are the most responsive of all.

Need for "Standard Stress" in Demonstrating Differences Between Patients and Controls

Another question which we sought to answer was whether a certain level of arousal must be reached in order to demonstrate differences between patients and nonpatients or whether such differences could be obtained under resting, "basal" conditions. From reviewing the literature prior to conducting our own experiments, we were led to suspect that some stimulation would be necessary because experiments which had been carried out under resting conditions had usually yielded negative or inconclusive results.

Our findings did indeed clearly show that, in differentiating between patients and controls, some form of stimulation was definitively superior to merely taking records under resting conditions. This has been demonstrated for blood pressure (**28**, p. 89), for muscle potentials in motor tasks (**31**, p. 54 and pp. 59 ff.), and again for muscle potentials in two separate investigations of startle (**30**, p. 327; **7**, p. 181). The only measure which we have found to discriminate well between patients and controls under "resting" conditions was frontalis-muscle potentials (**33**). However, we know that "resting" conditions associated with a testing session are by no means basal, and that—for example—significantly lower blood-pressure readings may be obtained from patients resting quietly on the ward than in the so-called "resting" condition of our experiments (**32**).

"Specific" vs. "nonspecific" stimulating situations. In producing higher levels of arousal in patients, is it necessary to present material to which patients are specifically sensitized or is it possible to demonstrate the difference between patients and controls by employing the same standard stimulating situation for all subjects? Our experiments clearly show that the latter is true. It is not necessary to present the patients with words or situations which have special meanings for them in order to produce more arousal in them than in controls.

As an example of a "specific-com-

plex" technique of producing high-level arousal, Luria (25) employed the method of controlled association in which he compared motor reaction to "critical" words (those which were especially arousing for the subject because of their association with specific life experiences) with reaction to indifferent words. Our situations, on the other hand, were chosen for their general arousal value, and we sought to avoid situations which would have special meaning for particular individuals.

With this point especially in mind we devised our standard situation of painful stimulation, because of the nearly universal avoidance reaction to pain. In order to permit more generalized conclusions, we also employed standard situations other than pain. One study is of especial interest because we reproduced the essential features of Luria's procedure, only substituting a series of size discriminations for the series of verbal stimulations which Luria employed (29). Conclusions from these experiments were as follows. All measures of motor activity recorded during performance of speeded size discrimination yielded reliable differences between patients and controls. In every instance there was evidence of greater physiological disturbance in the patients. The measures employed may be distinguished as skeletal-motor (motor control, muscular tension) and autonomic (systolic blood pressure). These differences in motor activity were manifested even though psychoneurotics, acute psychotics, and controls were practically identical with respect to perceptual performance.

These results led us to question certain views concerning determinants of higher arousal levels in psychoneurotics. In much current writing there is the underlying assumption that physiological disturbances in the psychoneurotic can be accounted for entirely in terms of situational explanations. These writers assume that there is no need to look for pathology in central and motor mechanisms, because they believe that amount of physiological disturbance is commensurate with the special significance which the situation has for the patient. Implied in this view is the assumption that only those stimuli which, through learning, have acquired special meaning for the patient have the power to produce an "abnormal" level of arousal. It assumes that the patient may participate in many situations without showing abnormally high levels of physiological reaction.

However, this view may well be questioned because it does not appear to fit with clinical observations. Cameron has written as follows:

> It will be noted that nearly all such patients [with anxiety states] complain that they cannot go into crowded places or into any situation where sustained efforts will be required of them. Their symptoms are made more severe by anything which elicits emotional reactions, such as altercations or participating in a discussion of illness. Nearly all find, at least at first, that their symptoms are increased by visiting their former places of employment or meeting fellow-workers. In other words, their symptoms are exacerbated by anything which serves to increase tension. *Emphasis should be placed upon the fact that their symptoms are elicited or intensified, not primarily by the reactivation of any conflict situation which may exist, but literally by everything in the course of the day which serves to increase tension* (5, pp. 56–57. Italics mine).

In therapy, relaxants of various kinds are devised to "damp" the "autonomous" reaction before proceeding with psychotherapy (41).

Strong auditory stimulation. Strong auditory stimulation served as another and very different kind of standard stimulating situation for comparing patients and controls. Two separate studies, the first one (30) with induced tension (produced by squeezing a rubber bulb), and the second (7) without induced tension and with a less intense stimulus, agreed in showing that the

most reliable difference between anxious patients and controls was in "after-response" following the period of primary reflex-startle reaction.

NEUROPHYSIOLOGICAL CONSIDERATIONS

In the interpretation of our findings in the experiments on strong auditory stimulation (7, 30), we cited the parallel between these observations on patients and findings in neurophysiological experiments on the reticular formation. In certain animal preparations, after-discharge in the cerebral cortex was abolished by stimulation in the reticular formation of thalamus (20) and brain stem (35). We believe that it is reasonable to suggest that some such inhibitory mechanism (as the one which abolished after-discharge) may be weakened in pathological anxiety.

Having implicated inhibition, we are required to examine this concept critically for a moment. Although there is by no means complete agreement on the matter of inhibitory mechanisms in the central nervous system, present evidence appears to point more and more in the direction of inhibition as a phenomenon in its own right, independent of excitation (i.e., not merely absence of excitation).

Of the current theories of inhibition known to me, Eccles' view seems most reasonable (12). Eccles and his coworkers developed a technique for placing a microelectrode within a single spinal motoneurone, and they were thus able to observe the electrical potential between the inside and the outside of the cell. They observed that when they stimulated an inhibitory nerve fiber it increased the polarization of the nerve cell on which it ended. Eccles called this effect "hyperpolarization," which, electrically is the opposite of what occurs when a nerve cell is fired (depolarized).

While Eccles' work was done on cells in the spinal cord, it nonetheless seems reasonable to suggest that the reticular formation could produce widespread inhibition in the cortex by hyperpolarizing cortical cells. Because the study of neuronal discharge in the cortex is a new field of research, sufficient data to decide this point are not at hand. But data which are presently available seem to be in line with the proposition that some impulse arriving in the cortex may have facilitatory effects, while others may produce opposite results (21, Fig. 19, p. 62).

If Eccles' theory is essentially correct,[5] we may work with inhibition as an independent process, and seek to understand the pathology of anxiety in terms of weakened inhibition. To make matters more concrete, we may draw on Eccles' hypothesis of a chemical transmitter for inhibition (12, p. 163) and on the recent experimental work of Elliott and Florey (13) to suggest that, in anxiety, the effectiveness of this substance has been reduced.

THE PROBLEM OF ETIOLOGY

The disorder of pathological anxiety may be conceived of almost entirely in terms of constitutional factors. It is logical to consider that certain individuals may inherit a deficient inhibitory mechanism. Such a person would consistently suffer from inability to relax throughout life, and would be seriously limited in the amount of stimula-

[5] Recent findings, although supporting Eccles' main conclusions, suggest that the phenomenon may be somewhat more complex than he originally supposed. The observations of Kuffler and Eyzaguirre (22) on inhibition of stretch receptor organs in crustaceans indicate that the polarity of the "inhibitory potential" varies with the state of the cell. When the cell is depolarized, an inhibitory volley causes polarization; when the cell is resting, an inhibitory volley causes depolarization.

tion that he could withstand. In such a case the constitutional weakness, rather than learning, would be the primary factor in etiology. While constitutional differences of genetic origin may account for degree of predisposition to the pathological condition of anxiety, clinical evidence stands against a purely genetic etiology. The fact that such a large number of patients recover from anxiety states (**17**, **34**) argues against a purely genetic-constitutional explanation of pathological anxiety.

Declining the genetic-constitutional explanation of anxiety implies that learning mechanisms are somehow involved in the pathology. In order to understand the full implications of this point of view, it is helpful to consider that degree of arousal is not a "given" in the stimulating situation. The same stimulating situation may produce quite different levels of physiological reaction in different persons, depending upon the effects of past learning.[6] We may compare individuals with respect to their physiological reactions in a large number of different situations. We may find, for example, that a certain person generally shows significantly higher levels of physiological reaction than most other individuals. If this person can avoid stimulating situations with high-arousal values he appears no different from others. However, in ordinary, everyday living, it is unlikely that he will be able to avoid such situations, and he will, therefore, be more or less constantly operating at physiological levels which are higher than normal. We may conjecture that in such a case in which stimulation keeps physiological levels constantly very high, over a long period of time there will be a weakening of inhibitory mechanisms from overuse.

[6] The reader will recall that in our physiological studies of psychiatric patients we attempted to avoid experimental situations which had special meanings for particular individuals. In an earlier section of this paper we referred to these situations as "nonspecific." We assume that an anxiety-prone individual, before he actually develops the pathological state (and after he recovers from it), will not show higher arousal levels in such "nonspecific" stimulating situations. The stimulating situations referred to as producing quite different levels of physiological reaction in different persons are, of course, what we called "specific" in the earlier section of this paper.

FURTHER CLINICAL-EXPERIMENTAL CONSIDERATIONS

Anxiety in combat. If our theory is correct, anxiety may be considered as a "disease of overarousal" (or in Selye's [**36**] terms, a disease of "adaptation"). That is, the critical neural change is thought of as being produced by a process of attrition from excessive and extended overarousal. It would not matter whether this overarousal were produced in an individual whose previous learning made him more prone to overarousal, or whether the individual were anxiety-resistant from past training, and was simply "overexposed" to situations (like battle) that everyone reacts to with extremely high physiological levels. With this view we can readily understand why under battle conditions each soldier would have his "breaking point," and why despite resistance to overarousal from constitution and previous learning, if situations of high-arousal level are repeated over a long enough time period, the critical change will finally occur. This seems to be the picture which emerges from studies of anxiety in combat (**16**, pp. 85 ff.).

Inhibitory Deficiency in Anxiety and in Manic States

From the clinical point of view, Cameron (**6**, p. 388) has drawn attention to the prominence of overactivity in the anxiety states. Cameron is inclined to believe, however, that the manic state best represents "pure overfacilitation," in comparison with anx-

iety, which he has described as "curbed overactivity." In drawing this comparison, Cameron was influenced by his careful observation of body movements. He found that the typical anxious patient was restless and in constant movement, but that he did not have the open, wide, flung-out movements of the manic. In general, the movements of the anxious patient remained within the body silhouette.[7]

The internally generated manic overactivity ("pure overfacilitation") could reasonably be accounted for by positing increased activity of facilitatory mechanisms.

PROBLEMS FOR FURTHER STUDY

The line of reasoning followed in the present paper suggests certain hypotheses which might be put to experimental test. In the first place, longitudinal physiological study of patients suffering severe states of anxiety should reveal changed physiological reaction under conditions of standard stimulation. That is, during performance of a motor task—for example, palmar skin conductance—electromyographic gradients and other physiological indicants of arousal should show decline when the patient is in remission, and should show increase again with relapse and return of the anxiety. This is a straightforward kind of investigation which one might suppose had already been undertaken. However, as far as I am aware, the study has not been carried out with anxiety states in the way proposed.

Anxiety and Learning

Physiological measures of arousal should prove valuable in learning experiments in which anxiety has been studied as a variable (**14**). For example, workers have employed questionnaires and scales (e.g., the Taylor scale [**40**]), to select subjects high in "anxiety." The chief purpose of such experiments has been to compare the learning speed of subjects scoring high on such a scale with other subjects scoring lower on the scale. It would appear that physiological measures could be applied to such problems with considerable advantage. Subjects who would probably react at high physiological levels could still be selected with the scales as an initial screening device; but physiological measurements could then be applied to provide actual values to place each subject on a continuum. Such methodology would appear promising in providing a continuous variable (i.e., physiological intensity, or arousal) for study in place of the rather dubious anxiety-nonanxiety dichotomy, and would have other advantages. For example, a low scorer on the scale might be temporarily upset, and so be misclassified in an experiment unless his actual physiological measures were available on the day of the experiment.

Research with Reserpine and Chlorpromazine

Patients exhibiting anxiety as the predominant symptom have been reported to improve significantly following the administration of reserpine and chlorpromazine (**19**). It should prove illuminating to study the effects of such drugs on physiological reaction of anxiety patients under controlled stimulating conditions. For example, with administra-

[7] On the surface, this appears incongruous with the notion of weakened inhibition. However, we may account for this constrained appearance of inhibition by suggesting the substitution of less efficient mechanisms of inhibition for the one which has suffered impairment. It may be, for example, that anxiety patients compensate for weakened autonomous mechanisms by calling on voluntary motor mechanisms (i.e., the pyramidal motor system). For example, in the absence of sufficient control from autonomous inhibitory mechanisms, the anxiety patient may avoid loss of motor control through co-contraction of antagonistic muscles.

tion of these drugs, would the electromyographic reaction of patients to strong auditory stimulation resemble the normal reaction more closely (show less after-response) than in the absence of the drugs?

It would likewise be of interest to determine the effect of such drugs on levels of physiological reaction in anxiety patients under conditions of moderate stimulation, such as those in our experiments with pain and with performance tasks. Would drug administration bring levels of physiological reaction down close to normal values under these conditions?

As a matter of fact, our experiments with psychiatric patients were performed prior to the full development of the concept of an intensity continuum in behavior, measured in terms of EMG gradients, level of palmar skin conductance, and other such physiological indicators. It would be highly desirable, therefore, to apply these more refined physiological measures to the study of anxiety patients. Do they, in fact, show steeper EMG gradients than normals in tracking, and are these gradients reduced in slope with administration of reserpine and chlorpromazine?

Proposed Animal Experiments

Certain aspects of these problems may be more advantageously studied with animal subjects. Studies of "experimental neuroses," as reviewed by Liddell (24), have shown that it is possible to produce chronic states characterized by physiological deviation. For present purposes it would be desirable to employ a form of stimulation which effectively maintains high levels of physiological reaction over long periods of time. For our purposes it would not matter particularly how the stimulation was produced; the main requisite is that high physiological levels be recorded continuously over days and weeks.

The main purpose of such an experiment would be to determine whether keeping physiological levels constantly high would finally produce "anxiety" in animals (i.e., animals with raised physiological levels in standard test situations). If such experiments did turn out positively, valuable animal "preparations" would be available for neurophysiological and pharmacological studies.

Such a "preparation" might be used, for example, to determine whether inhibitory effects from stimulation in the reticular formation are weaker than in normal animals. We might even conceive of an experiment paralleling the ones which we carried out with human subjects. It would seem possible to implant electrodes in the reticular formation to search for areas which fire inhibitory impulses to the cerebral cortex following strong auditory stimulation. Furthermore, pharmacological investigation (13) might be directed to the question whether there is an inhibitory substance in the brain which becomes dilute with long-continued overarousal.

SUMMARY

The main purpose of this paper is to consider some recent experimental data which suggest a way out of the present confusion surrounding the concepts of motivation, emotion, and anxiety. Two lines of investigation, each employing physiological methods, are examined. In one experimental program, measures such as steepness of muscle-potential gradients and level of palmar skin conductance were found to be useful indicants of arousal level. The results of several experiments demonstrated significant relationships between such physiological indicants and excellence of performance on various motor tasks, such as mirror tracing and tracking. In this empirical setting, problems of relationship between concepts of motivation and emotion are reconsidered.

The arousal concept is then applied to the problem of pathological anxiety in psychiatric patients. The earlier results from physiological studies carried out with psychiatric patients as subjects are reviewed in the light of the more recent work on physiological indicants of arousal. Considerable confusion has arisen because the term "anxiety" has been used to denote two quite different states of the organism: (*a*) any increase in level of arousal, however brief the rise (or however selective the stimulating condition); and (*b*) a pathological state in which the patient appears chronically overreactive (physiologically) to every stimulating situation.

It seems reasonable to restrict the term "anxiety" to the chronic pathological condition. Results from physiological studies carried out with patients suffering this pathological condition indicated that standard stimulation (or "stress") accentuated the differences in arousal between anxiety patients and controls. Under resting conditions such differences were usually insignificant. On the basis of the data reviewed, certain hypotheses concerning the nature and etiology of pathological anxiety are tentatively advanced. It is suggested that anxiety may be produced in an individual (in animal as well as in man) by keeping level of arousal very high over long periods of time. Finally, recent neurophysiological findings are cited in stating the hypothesis that such continuous overarousal may result in impairment of central inhibitory mechanisms.

REFERENCES

1. BARTOSHUK, A. K. Electromyographic gradients in goal-directed activity. *Canad. J. Psychol.*, 1955, **9**, 21–28.
2. BARTOSHUK, A. K. Electromyographic gradients as indicants of motivation. *Canad. J. Psychol.*, 1955, **9**, 215–230.
3. BÉLANGER, D. J. "Gradients" musculaires et processus mentaux supérieurs. *Canad. J. Psychol.* (in press).
4. BORING, E. G. *The physical dimensions of consciousness.* New York: Century, 1933.
5. CAMERON, D. E. Autonomy in anxiety. *Psychiat. Quart.*, 1944, **18**, 53–60.
6. CAMERON, D. E. Some relationships between excitement, depression, and anxiety. *Amer. J. Psychiat.*, 1945, **102**, 385–394.
7. DAVIS, J. F., MALMO, R. B., & SHAGASS, C. Electromyographic reaction to strong auditory stimulation in psychiatric patients. *Canad. J. Psychol.*, 1954, **8**, 177–186.
8. DELAFRESNAYE, J. F. (Ed.). *Brain mechanisms and consciousness.* Springfield, Ill.: Thomas, 1954. (See especially discussion by H. H. Jasper, p. 132.)
9. DUFFY, ELIZABETH. The conceptual categories of psychology: a suggestion for revision. *Psychol. Rev.*, 1941, **48**, 177–203.
10. DUFFY, ELIZABETH. A systematic framework for the description of personality. *J. abnorm. soc. Psychol.*, 1949, **44**, 175–190.
11. DUFFY, ELIZABETH. The concept of energy mobilization. *Psychol. Rev.*, 1951, **58**, 30–40.
12. ECCLES, J. C. *The neurophysiological basis of mind.* Oxford: Clarendon, 1953.
13. ELLIOTT, K. A. C., & FLOREY, E. Factor I—Inhibitory factor from brain. Assay. Condition in brain. Simulating and antagonizing substances. *J. Neurochem.*, 1956, **1**, 181–192.
14. FARBER, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation.* Lincoln: Univer. of Nebraska Press, 1954.
15. FREEMAN, G. L. *The energetics of human behavior.* Ithaca, N. Y.: Cornell Univer. Press, 1948.
16. GRINKER, R. R., & SPIEGEL, J. P. *Men under stress.* Philadelphia: Blakiston, 1945.
17. HARRIS, A. The prognosis of anxiety states. *Brit. med. J.*, 1938, **2**, 649–664.
18. HEBB, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.
19. HOLLISTER, L. E., TRAUB, L., & BECKMAN, W. G. Psychiatric use of reserpine and chlorpromazine. Results of double-blind studies. In N. S. Kline (Ed.), *Psychopharmacology.* Washington, D. C.: Amer. Assoc. for Advancement of Science, 1956.
20. JASPER, H. H. Diffuse projection systems: the integrative action of the thalamic

reticular system. *EEG Clin. Neurophysiol.*, 1949, **1**, 405–420.

21. JUNG, R. Neuronal discharge. *EEG Clin. Neurophysiol.*, 1953, Suppl. No. 4, 57–71.
22. KUFFLER, S. W., & EYZAGUIRRE, C. Synaptic inhibition in an isolated nerve cell. *J. gen. Physiol.*, 1955, **39**, 155–184.
23. LACEY, J. I. Individual differences in somatic response patterns. *J. comp. physiol. Psychol.*, 1950, **43**, 338–350.
24. LIDDELL, H. S. Conditioned reflex method and experimental neurosis. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944. Vol. I, pp. 389–412.
25. LURIA, A. R. *The nature of human conflict.* New York: Liveright, 1932.
26. MALMO, R. B., & DAVIS, J. F. Physiological gradients as indicants of "arousal" in mirror tracing. *Canad. J. Psychol.*, 1956, **10**, 231–238.
27. MALMO, R. B., & SHAGASS, C. Physiologic studies of reaction to stress in anxiety and early schizophrenia. *Psychosom. Med.*, 1949, **11**, 9–24.
28. MALMO, R. B., & SHAGASS, C. Studies of blood pressure in psychiatric patients under stress. *Psychosom. Med.*, 1952, **14**, 82–93.
29. MALMO, R. B., SHAGASS, C., BÉLANGER, D. J., & SMITH, A. A. Motor control in psychiatric patients under experimental stress. *J. abnorm. soc. Psychol.*, 1951, **46**, 539–547.
30. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. A method for the investigation of somatic response mechanisms in psychoneurosis. *Science*, 1950, **112**, 325–328.
31. MALMO, R. B., SHAGASS, C., & DAVIS, J. F. Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.*, 1951, **12**, 45–66.
32. MALMO, R. B., SHAGASS, C., & HESLAM, R. M. Blood pressure response to repeated brief stress in psychoneurosis: a study of adaptation. *Canad. J. Psychol.*, 1951, **5**, 167–179.
33. MALMO, R. B., & SMITH, A. A. Forehead tension and motor irregularities in psychoneurotic patients under stress. *J. Personality*, 1955, **23**, 391–406.
34. MILES, H. H. W., BARRABEE, EDNA L., & FINESINGER, J. E. Evaluation of psychotherapy. *Psychosom. Med.*, 1951, **13**, 83–105.
35. MORUZZI, G., & MAGOUN, H. W. Brain stem reticular formation and activation of the EEG. *EEG Clin. Neurophysiol.*, 1949, **1**, 455–473.
36. SELYE, H. *Stress.* Montreal: Acta, 1950.
37. STENNETT, R. G. The arousal continuum. *J. exp. Psychol.* (in press).
38. STENNETT, R. G. The relationship of alpha amplitude to the level of palmar conductance. *EEG Clin. Neurophysiol.*, 1957, **9**, 131–138.
39. SURWILLO, W. W. Psychological factors in muscle-action potentials: EMG gradients. *J. exp. Psychol.*, 1956, **52**, 263–272.
40. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.
41. TYHURST, J. S., & RICHMAN, A. Clinical experience with psychiatric patients on reserpine—preliminary impressions. *Canad. med. Assoc. J.*, 1955, **72**, 458–459.
42. WALLERSTEIN, H. An electromyographic study of attentive listening. *Canad. J. Psychol.*, 1954, **8**, 228–238.

(Received December 10, 1956)

Psychological Review
Vol. 64, No. 5, 1957

OPTICAL MOTIONS AND TRANSFORMATIONS AS STIMULI FOR VISUAL PERCEPTION[1]

JAMES J. GIBSON

Cornell University

How do we see the motions of objects in the world around us? The way to go about answering this question is to note the kinds of physical motion that occur in the human environment and then to examine the kinds and variables of optical stimulation that correspondingly occur. The isolation and control of these variables with suitable optical apparatus will make possible an experimental psychophysics of kinetic impressions. The desirability of such a psychophysical approach has been pointed out in an earlier paper (4) and the following proposals modify or make explicit a number of suggestions there made.

Distinction Between Physical Motions and Optical Motions

Ever since Isaac Newton supposed that the motions of things revealed the forces behind them and thereby the causes of all events, physics has been concerned with the observation of motions. The beauty of the idea for physics is that it applies to *all* things: stars and planets; stones, machines, and animals; particles and atoms. Of these motions, however, only a certain class is the concern of perceptual psychology. The things whose motion is visible are substances which, in the first place, differentially reflect or emit light and, in the second place, are not either too far away or too small. The motion of the wind is invisible because gaseous substances transmit light instead of reflecting it. The motion of the heavenly bodies is invisible because their angular change of position is too small per unit of time. And the motion of microscopic bodies is invisible because their boundaries reflect an optical texture too small for the eye to resolve. But the environment of man contains an enormous variety of surfaces which do project focusable light to his eye, and these are the bodies the psychologist must be interested in. For when they fall, rise, turn, roll, bend, flow, twist, writhe, stretch, or break, an eye can register this event and the animal possessing the eye can respond to it.

We may observe that physical motions can be classified as *rigid* or *nonrigid*, the former being characteristic of crystalline substances and the latter of elastic or fluid substances. Rigid physical motions are exemplified in mechanics; they are analyzable into components of translation and rotation on or around any of three axes, and they have been studied since the time of Newton. Nonrigid physical motions are exemplified in biology since the growth and also the reactive movements of living animals are generally of this sort (10).

The motions of the physical environment might also be classified as *connected* or *nonconnected*. In the former the parts of the moving substance remain adjacent, even if not rigid, whereas in the latter they do not and are not even considered parts of the "same"

[1] This essay is part of an address entitled *Stimulation and Perception* delivered by the author in September 1955 at San Francisco as retiring president of the Division of Experimental Psychology, APA. The film which accompanied it, and serves to illustrate this paper (5), was produced with the support of the Office of Naval Research under Contract NONR 401(14) with Cornell University, which is concerned with research on the perception of motion and space.

substance. Instead they are treated as separate motions of different objects. The separations and fusions, attractions and repulsions, or collisions of things, animals, and people are all of this sort.

Physical motions are given to an eye only in the form of optical motions. An optical motion is an event in the *optic array*, that is, in the light reflected from an illuminated environment to an eye or, rather, to any position in the air where an eye might be placed (7). An eye is an organ for exploring an optic array. The solid sector of this array which an eye takes in is the basis of patterned vision; neither objects nor their motions could affect the eye except by means of it. External motion can be seen only if some differentiated part of the array is displaced relative to the rest of it, or to some other part, or if parts move relative to one another. There has to be some change of its pattern, considered as a projection to a point.

An optical motion, then, is a projection in two dimensions of a physical motion in three dimensions. The projection may be taken either as on the surface of a sphere centered at the eye or as on a plane in front of the eye. When locomotor movements of the observer are to be considered, the former is preferable (8), but when, as here, they are not involved, the plane projection is better. The one can be converted into the other if necessary. Our question is, What kinds of optical motion occur which might serve as stimuli for perception?

KINDS AND VARIABLES OF OPTICAL MOTION

How can optical motions be described or specified? The question has to do with the motions of a texture or pattern in a two-dimensional array. Tentatively, there seem to be two general possibilities. First, one could divide the pattern into convenient elements, describe the positions of all the elements by pairs of coordinate values, and finally describe the motions of all the elements by the successive pairs of values. Or one could describe the motions of the elements by direction and speed at successive moments of time. This procedure is analytical. Second, a nonanalytical method of specifying optical motions is possible. One could simply take the pattern as given, and then use the operation defining a *perspective transformation* in geometry to describe a family of changes of pattern. This method does not exhaust all the varieties of optical motion, as will be evident, but it has advantages for an experimenter who needs to produce an artificial optical array.

Continuous perspective transformations. In geometry, any form or pattern on one plane which is a projection of a form or pattern lying on some other plane is called a *perspective transformation.* These forms are static. When the point of projection (the focus of the sheaf of lines which connect the pair of forms, point for point) is near the two planes, we speak of a *central* or *polar* projection; when this point of projection is at a very great distance from the planes, we speak of a *parallel* projection. (It may be useful to recall that the "plans" and "elevations" of an architect are cases of parallel projection, but that his drawing for the client's eyes is a case of polar projection. This latter is the case we are chiefly concerned with.) When the two planes are parallel, the difference between the projected form and the given form is one of size only; it is called a *similarity* transformation. When the two planes are not parallel, the difference between the forms is that to which common meaning applies the term "perspective," or sometimes the term "foreshortening." In geometry,

both the difference of size and that of shape are classed as perspective transformations.

We can now speak of motion. When the angle or the distance of the first plane relative to the second plane is altered, the projected form is correspondingly altered. The fact is that the relation between any earlier and any later projected form is *also* a perspective transformation. The relation holds between any two of its stages in times. Hence, the motion in question may be described as a continuous series of perspective transformations. It is a relation between a temporal series or family of static forms, not merely between two forms. Any such moving transformation can be analyzed by six parameters corresponding to the six components of the possible movements of the first plane—that is, three of translation and three of rotation.

Families of continuous perspective transformations. It can now be observed that all optical motions resulting from the rigid physical motions of the flat face of an external object are continuous families of perspective transformations, as defined above. These are optical motions as taken on a plane in front of the eye. The six parameters of optical motion can be visualized as (*a*) vertical translation of the pattern in the plane, (*b*) horizontal translation of the pattern, (*c*) enlargement or reduction of the pattern, (*d*) horizontal foreshortening of the pattern, (*e*) vertical foreshortening of the pattern, and (*f*) rotation of the pattern in the plane.

These parameters of transformation are for the case of a *polar* projection. As the focus of projection is taken at an increasing distance from the two planes, one approaches the case of a *parallel* projection. For the latter case, three of the six parameters have been altered in character, namely (*c*), (*d*), and (*e*) above, while (*a*), (*b*), and (*f*) remain unaltered. Enlargement or reduction of the pattern has vanished; horizontal foreshortening becomes a mere horizontal flattening; and vertical foreshortening becomes vertical flattening.

These "pure types" of optical motion can be observed on a motion picture screen (**5**). An irregular or regular contour shape or an irregular or regular group of spots can be made to undergo continuous perspective transformations, and the observer can note the various types of motion in the plane of the screen. One can note, for example, that in types (*d*) and (*e*) a square is transformed into a trapezium with polar projection, but is transformed into a flattened rectangle with parallel projection. The interesting fact, however, is that for types (*c*), (*d*), and (*e*) with polar projection it is *very difficult to notice* the motion in the plane. Instead, one sees a sort of "virtual" object or surface which (*c*) moves toward or away from the screen, (*d*) rotates on its vertical axis, or (*e*) rotates on its horizontal axis. One sees, in other words, rigid motion in depth. The suggestion is that the parameters of transformation are stimuli for the phenomenal parameters of the motion in space of one face of an object.

The rotations in depth are similar in some respects to the kinetic depth effect obtained by Wallach (**14**). Such effects have been observed for Lissajous figures (**3**), and long ago for the silhouette of a rotating windmill against the horizon (**1**, p. 270). All these apparent rotations are said to be characterized by ambiguity as to the direction of rotation, and by spontaneous reversals in the direction of rotation. The apparent rotations shown in the film, however, are *not* characterized by ambiguity or reversals of direction when the transformations are those obtained with polar

projection, but only when the transformations are those obtained with parallel projection.

A psychophysical experiment has been performed on the degree of perceived semirotation in depth as a function of the transformation sequence (6).

APPARATUS FOR PRODUCING CONTINUOUS PERSPECTIVE TRANSFORMATIONS IN THE OPTIC ARRAY

The method used to display the geometrical transformations on a motion picture screen was not by "animation" of film; the procedure, rather, was to photograph the window of a device which might be called a shadow transformer. Details of its construction are given in the report of the experiment (6). It will here be described only as it suggests possibilities for a psychophysics of motion perception. It consists of a translucent screen with a point source of light on one side and an observer symmetrically on the other side. In the square luminous window, dark shapes, patterns, or textures can be made to appear. They are shadows, not objects, so that only two grades of intensity exist, surface texture and binocular disparity are eliminated, and accommodation and convergence are controlled for this sector of the optic array to the eye. The variables of form and transformation are thus isolated for study.

In the diverging ray sheaf from the point source to the translucent window a mount can be placed, a pane of transparent material large enough so that it can be moved without its edges being projected on the screen. Forms, patterns, or textures cut from gummed paper or masking tape can be attached to this mount, or it can be traced with ink or even sprinkled with talcum powder, so that shadows of many varieties are projected on the screen. The mount can be rotated on any of three axes, or translated in any of three dimensions. Hence, considering the mount and the screen as two geometrical planes, changes in the position of the mount will yield all possible perspective transformations of the shadow relative to the shadow caster, and likewise all parameters of continuous perspective transformation of the shadow itself. Previous shadow-casting devices, most recently Wallach's (14), have not been constructed for this systematic purpose. They have also not utilized polar projection. For purposes of comparison, the present apparatus can also be illuminated by a projector

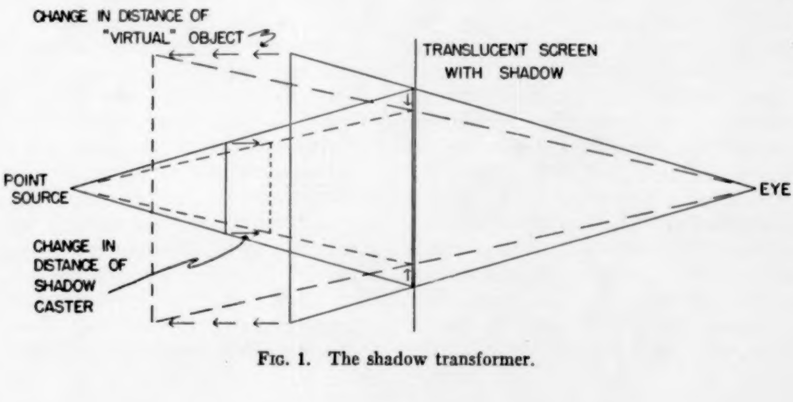

FIG. 1. The shadow transformer.

beam at 80 ft. instead of a point source at 5 ft.; the former yields an approximately parallel projection. The use of a transparent mount obviates the necessity of any visible support for the shadow caster, or of having the shadow it casts extend below the bottom of the window.

The optical geometry of the apparatus is given in Fig. 1 for a size transformation. It may be noted that the converging ray sheaf to the eye is the geometrical opposite of the diverging ray sheaf. The relation of the shadow to its "virtual object" (by analogy with the "virtual image" in a mirror) is simply the reverse of the relation of the shadow to the shadow caster. This reversal does not affect the transformations in any other way. It is true that the motions perceived are opposite to those of the shadow caster, but this does not present a paradox if one remembers that vision depends on the optic array, not on the way an optic array can be artificially produced.

INTERNAL DEPTH OF TRANSPARENT OBJECTS AND COLLECTIONS OF OBJECTS

So far, consideration has been limited to continuous transformations of textures corresponding to the opaque face of an object. What about the optical motions corresponding to the physical motions of transparent objects, or things with different parts in depth, or rigid collections of things like a forest of trees or a tangle of wire? The phenomenon termed "stereokinesis" (**11**, Ch. 13 and references) and the kinetic depth effect of Wallach (**14**) involve an impression of *internal depth*. The "virtual object" of these experiments is often a collection of posts or a figure of bent wire.

The complex optical motions of these experiments can be analyzed by considering different parameter values of the same perspective transformation corresponding to the different planes of depth of the object. This suggestion becomes clearer with a concrete example. If, with the shadow transformer, one casts a shadow on the screen through several parallel sheets of glass, each of which has been sprinkled with talcum powder, the composite shadow looks something like the Milky Way in the night sky. When the sandwich of mounts is *moved*, however, the perception separates into clear planes of depth, each layer of nebulous material standing in front of its neighbor. The effect is shown in the film (**5**).

The internal depth of the virtual object produced by the moving shadow of a bent wire, or by an arrangement of vertical sticks on a horizontal turntable, has been frequently studied. The writer suspects that it will yield to the kind of analysis proposed. The depth is similar to binocular stereoscopic depth; a unifying hypothesis would be that the simultaneous disparity of binocular images is only a special geometrical case of the successive disparities of a continuous image. Both rest on the geometry of parallax, that is, the projection of a collection of objects in space to a point in space, and this has led Tschermak to include both under the term "parallactoscopy" (**13**). Both are transformations in the most general sense of the term, and perhaps both sorts of disparity should be treated as transformations.

The impression of surfaces meeting at a corner. With the apparatus described, one can also produce the impression of *more* than one flat face of an object moving in depth. If two transparent mounts are joined to make an angle, and if both are given some opaque texture of any kind, the combined shadow looks like a plane surface so long as the combined mounts are kept stationary. When they are moved, however, the shadow

becomes two faces or surfaces making a corner. The relative slant of one surface to the other can be judged with some accuracy.

WHAT IS THE STIMULATING EFFECT OF NONPERSPECTIVE TRANSFORMATIONS IN THE OPTIC ARRAY?

If the rigid mechanical motions of the physical environment are represented by one kind of geometrical transformations in the pattern of light, are the nonrigid biological motions of the environment represented by a *different* kind of geometrical transformations in the pattern of light? The difference is suggested when the geometer describes topology as "rubber sheet geometry." This is concerned with changes of bidimensional form *other* than the changes heretofore described. Considering an organism in silhouette, its reactive movements cannot be compounded of the six pure types of optical motion considered above. Neither can its growth be described as magnification. Medawar, a biologist following D'Arcy Thompson, seems to have demonstrated that the change of shape of the human figure from infancy to adulthood, disregarding the change of size, is a specific continuous transformation which can only be suggested in words. A "tapered stretch" describes it approximately (10, pp. 177 ff.). The change is monotonic, i.e., it keeps the same trend. Geometrically, there are different *forms* of change of bidimensional form. Conceivably, the visual mechanism is sensitive to these forms of change.

The shadow transformer can be adapted to display nonperspective transformations if an elastic or flexible sheet is used for the transparent mount which carries the shadow-casting form or texture, and if this is stretched or bent in some way. Preliminary observations suggest that the resulting perception of motion is correspondingly elastic instead of rigid. There are technical difficulties in controlling such optical motions. But if apparatus can be built for systematically producing them, it will be open to the perceptual psychologist to study such phenomena as animate, expressive, and physiognomic movements in the manner of psychophysics.

DISJUNCTIVE OR SEPARATE OPTICAL MOTIONS

The converting of a single form on a plane into *two* forms is something which goes beyond the continuous transformations heretofore considered. There is, instead, a discontinuity in both the temporal and the spatial series. The geometer is tempted to describe it by saying that there is a breaking or tearing of the surface, thus falling back on a physical analogy.

If certain parts of a connected optical pattern undergo one kind of transformation while other parts undergo another kind, it might be predicted that the perceived surface will become two perceived surfaces, each composed of the parts carrying the same transformation. This is obvious when one set of parts moves in one direction and the other in a different direction, and the fact was recognized in Wertheimer's law of "common fate" as a determiner of sensory organization (11, Ch. 13). It should equally be true, however, when one set of parts carries a slant transformation or a size transformation different from that of the other; the texture will break into two textures each moving in its own tridimensional way. Some of these possibilities have been investigated by using a "sandwich" of mounts in the apparatus, and the film illustrates a few of these possible dual disjunctive motions. Perceptual separation does result. Evidently when the parts of an optical texture undergo *joint motion* this does not

have to be understood as a set of motions with the same velocity in the same direction.

It is also possible to note what happens when *all* the parts of a connected optical pattern move, each in a different direction: the pattern becomes many smaller objects, like a swarm of ants. This result also suggests that what connected the elements of the pattern in the first place was their nonmotion relative to one another; in the optic array, after all, stability is only a special case of transformation. Research on the problem of how elementary motions might be *organized* into a single unitary motion has been performed by Johannson (9), Duncker (2), and Metzger (11).

The varieties and dimensions of optical motion in which the parts are *not* connected in adjacent order are of formidable complexity. It is not even clear how to go about classifying them. Disjunctive optical motions are, however, the stimuli by virtue of which we see occurrences, happenings, and actions in the world around us. There is certainly order and lawfulness in them, for Michotte has studied the impressions of causality induced by higher order variables of nothing more objective than the motions of separate spots (12). These abstract variables are clearly discriminable by observers, for the impressions can be made to come or go as the experimenter varies certain spatiotemporal conditions.

CONCLUSIONS

If the optical geometry here expounded is correct, there is a possible basis in optical stimulation for the ability to distinguish between and among rigid, elastic, and multiple moving things. The basis lies in different mathematical modes of transformation and motion. The implication is that we see both the constant and the changing properties of things in this way. We see them not because we have formed associations between the optical elements, not even because the brain has organized the optical elements, but because the retinal mosaic is sensitive to transformations as such. These are stimuli for perception.

Is it really plausible, one might ask, to call anything as apparently abstruse as a continuous series of transformations a *stimulus?* A bit of evidence may here be convincing. A puff of air to the cornea of the eye is a stimulus for the blink reflex in the pure and original sense of the term. The fact is that when an observer with the apparatus described is near the screen, a rapid expansion of the shadow until it fills the screen will also produce a blink reflex. The latter event ought to be considered as much a stimulus as the former.

REFERENCES

1. BORING, E. G. *Sensation and perception in the history of experimental psychology.* New York: D. Appleton-Century, 1942.
2. DUNCKER, K. Über induzierte Bewegung. *Psychol. Forsch.,* 1929, **12,** 180–259.
3. FISICHELLI, V. R. Effect of rotational axis and dimensional variations on the reversals of apparent movement in Lissajous figures. *Amer. J. Psychol.,* 1946, **59,** 669–675.
4. GIBSON, J. J. The visual perception of objective motion and subjective movement. *Psychol. Rev.,* 1954, **61,** 304–314.
5. GIBSON, J. J. *Optical motions and transformations as stimuli for visual perception.* Motion picture film. Psychological Cinema Register, State College, Pa., 1955.
6. GIBSON, J. J., & GIBSON, ELEANOR J. Continuous perspective transformations and the perception of rigid motion. *J. exp. Psychol.,* 1957, **54,** 129–138.
7. GIBSON, J. J., PURDY, J., & LAWRENCE, L. A method of controlling stimulation for the study of space perception: the optical tunnel. *J. exp. Psychol.,* 1955, **50,** 1–14.
8. GIBSON, J. J., OLUM, P., & ROSENBLATT, F. Parallax and perspective during aircraft

landings. *Amer. J. Psychol.*, 1955, **68**, 372–385.

9. JOHANNSON, G. *Configurations in event perception.* Uppsala: Almqvist and Wiksell, 1950.
10. LE GROS CLARK, W. E., & MEDAWAR, P. B. *Essays on growth and form presented to D'Arcy Thompson.* New York: Oxford Univer. Press, 1945.
11. METZGER, W. *Gesetze des Sehens.* Frankfurt: Waldemar Kramer, 1953.
12. MICHOTTE, A. *La perception de la causalité.* Louvain: Publications Universitaires, 1954.
13. TSCHERMAK-SEYSENEGG, A. Über Parallaktoskopie. *Arch. f. d. ges. Physiol.*, 1939, **241**, 454–469.
14. WALLACH, H., & O'CONNELL, D. N. The kinetic depth effect. *J. exp. Psychol.*, 1953, **45**, 205–217.

(Received July 18, 1956)

Psychological Review
Vol. 64, No. 5, 1957

SOME PROBLEMS OF ECLECTICISM

MARY HENLE

Graduate Faculty of Political and Social Science, New School for Social Research

Some ten years ago Woodworth, commenting on the situation in psychology as a whole, wrote: "Some may lean toward one school and some toward another, but on the whole the psychologists of the present time are proceeding on their way in the middle of the road . . ." (**21**, p. 254). He suggested that "If we could assemble all these psychologists [all the psychologists in the world] in a convention hall and ask the members of each school to stand and show themselves, a very large proportion of the entire group would remain seated" (**21**, pp. 254–255).

A similar position was taken by Boring at about the same time:

> During the 1930's the *isms* pretty well dropped out of psychology. . . . The only reason for mentioning these four schools in this book is that the student hears about *behaviorism* and *Gestalt psychology* and has a right to be told what they are and that they are no longer important as schools. What was good in all the schools is now simply part of psychology (**4**, p. 11).

The eclectics, rising above the conflict of schools, hold that psychologists today are in happy agreement. It is their position that no real issues exist among the various points of view in psychology. For it is only on the basis of such a belief that one is able to select from each approach, combining the theoretical contributions of all. As Woodworth presents the middle-of-the-road position:

> Every school is good, though no one is good enough. . . . One points to one alluring prospect, another to another. . . . Their negative pronouncements we can discount while we accept their positive contributions to psychology as a whole (**21**, p. 255).

If eclecticism is as prevalent in contemporary psychology as the above statements suggest, it becomes important to examine its consequences for theory. It is the hypothesis of the present paper that the eclectics have, to a large extent, succeeded in reconciling differences only by obscuring theoretical issues. An alternative to this kind of eclecticism will be proposed.

We may begin with an examination of specific instances of eclectic reconciliations of differences. Examples of two kinds of eclecticism will be discussed, one having to do with reconciliations of positions which refer to the entire field of psychology—the attempt to resolve conflicts among "schools" or general points of view—and a more circumscribed eclecticism relating to particular psychological problems.

As an instance of the first kind, Woodworth [1] writes:

> A broadly defined functional psychology starts with the question "What man *does*" and proceeds to the questions "How?" and "Why?" . . . So broadly defined . . . functional psychology scarcely deserves the name of a school because it would include so many psychologists who have not professed themselves. Now the question is whether our middle-of-the-roaders are not after all members of this broadly conceived functional school. . . . But if the middle-of-the-roaders are really functionalists, the question is then whether the same would not be true of all the schools. Are they not all functionalists at heart? (**21**, p. 255).

Commenting on such a functionalism, Boring wrote in 1950: "Woodworth be-

[1] A number of the examples to be considered will be taken from Woodworth's writings because the present author regards him as one of the clearest of the eclectics and one of those whose theories are to be taken most seriously.

lieved that psychologists were more in agreement than their quarrels indicated, and he sought a system to which all could subscribe. He very nearly succeeded (3, p. 565).

It is not likely to be denied that psychology today has a functionalist flavor. The interest in the adaptive value of psychological processes is everywhere apparent. Nor can there be much disagreement with a functionalism defined in terms of these three questions of Woodworth's. But it tells us very little about a psychologist to say that he is a functionalist in this sense. What we need to know are the kinds of answers which a particular psychology gives to these questions. These are the issues in contemporary psychology, and here it is that disagreements arise. What, for example, is the relation of reinforcement, or of repetition, to learning? How does the learning process proceed? What is the nature of the fundamental human motives? How does the group exert its influence on the individual? It is the answers to such questions as these that divide psychologists. It would seem that Woodworth has succeeded in bringing all together only by obscuring such issues. If all are functionalists today, we still have the problem of examining the differences among the several varieties of functionalism.

The point may be made more specifically in connection with the same author's theory of the conditions of transfer of training, the doctrine of identical components (**19**), reformulated by Woodworth and Schlosberg (**23**) as a theory of "common factors." It is put forward to resolve the differences between those who hold that transfer is a function of the identical elements in two learning tasks, and those who maintain that it is a matter of the application of common principles or other whole properties to the two activities. Woodworth's view [2] is that anything concrete can be transferred—thus both identical elements and principles—since "any idea that can be recalled, or any attitude that can be reinstated is concrete enough to qualify. Perhaps anything that can be learned can be transferred" (**19**, p. 207). Again, "what is successfully transferred is usually something you can put your finger on—a principle, a good emotional attitude, a technique" (**22**, p. 582).

The controversy about the conditions of transfer is settled, in other words, by saying that *something* is carried over from one activity to the other. It is true that this is a formulation which covers most of the cases. But it lacks an advantage of both of the theories it displaces, namely the attempt to state the specific conditions of transfer. It would seem that the differences are resolved only at the expense of any specific theory in the area of the controversy. The theory covers all the cases only by telling us nothing specific about any of them.

To return to the more general kind of eclecticism, there exist today a number of efforts to reconcile the various significant theoretical positions in psychology. A number of authors have attempted to resolve the differences between behavior theory and psychoanalysis, Gestalt psychology and psychoanalysis, behavior theory and Gestalt theory.[3] Several examples of this trend in contemporary psychology will be examined.

[2] Since the earlier formulation (**19**) is the more explicit, and since the later (**23**) seems not to differ from it in principle, the former will be drawn upon here.

[3] Not all of these discussions are eclectic, at least in the meaning used here. Some attempt to understand the contributions of one psychology in terms of the theory of another, rather than to resolve differences between them. Such work, for example *Personality and Psychotherapy* by John Dollard and Neal E. Miller, will not be considered here. It presents interesting problems of its own which deserve separate treatment.

The problems seen above in Woodworth's formulations exist also in Abt's statements about basic agreements between psychoanalysis and Gestalt psychology with respect to the structure and development of the personality:

Freud's multiple-structured self is not essentially different conceptually from Lewin's division of the person into regions. The dynamic and economic interchanges that are postulated as occurring with respect to the id, ego and superego in psychoanalysis find parallel expression in Lewin's system of barriers and the classes of movements across them (**1**, pp. 38–39).

If Abt means that Lewin's *metatheory* is compatible with Freud's *theory* of personality structure, a case could be (but has not been) made for this position. But to equate Lewin's division of the personality into regions with Freud's topographical analysis of the person is to lose all the specific psychological insights of the latter and much of the metatheoretical contribution of the former. Lewin's inner-personal regions, if translated into Freudian terms, would undoubtedly fail to distinguish between id, ego, and superego; his motor-perceptual region includes some but not all of the functions of the Freudian ego.[4] Lewin is, indeed, largely unconcerned with the specific content of the personality, with the distinction between conscious and unconscious motivation, and with the historical development of the person; it is impossible to discuss Freud's topographical divisions apart from such considerations. Again, Lewin has not discussed the content and nature of the forces responsible for behavior; thus the parallel with Freud's statements about dynamics can be maintained only if one turns one's attention away from the specifics of Freud's successive instinct theories. It is indeed possible to find parallels between Lewin's statements about the tendency of systems under tension to seek discharge and Freud's formulations about the pleasure principle; but to equate the two is to lose the specific character of both the pleasure and reality principles and to neglect Freud's theories about behavior which is independent of the pleasure principle (**8**).

A further illustration will be given of the tendency prevailing in contemporary psychology to reconcile the ideas of Freud and Lewin. Another author writes: "It is the thesis of this paper that a synthesis of the ideas of Lewin and Freud provides a basis for the beginnings of an integrated system of psychological theory . . ." (**5**, p. 206). More specifically, "We have already noted the structural parallelism between Freud's divisions of the personality and Lewin's psychical systems" (**5**, p. 222). This writer, it is true, limits the parallelism, remarking, "Freud has provided the living clay for the Lewinian scaffolding" (**5**, p. 228).[5] Still, closer examination suggests that essential differences between the "living clay" and the "scaffolding" have been neglected. For example, "One of these [defense mechanisms], projection, becomes the equivalent for Lewin's unreality" (**5**, p. 222). While there are, of course, important components of unreality in projections, the two concepts are by no means equivalent, if only because not all events on a level of unreality (for example,

[4] For example: "It is to a certain degree arbitrary where one draws the boundary between the motor-perceptual system and the inner regions, whether for instance one considers the understanding of speech as an event within the boundary zone or within the inner-personal systems" (**16**, p. 178).

[5] In this connection Bronfenbrenner comments on Lewin's neglect of the content of psychical systems: "This is indeed an unfortunate oversight" (**5**, p. 214). The thesis will be developed elsewhere that this is no oversight, but that Lewin undertook a different task.

dreams, fantasies, vague hopes and wishes) can be described as projections, at least as Freud uses the term. Again, to put the Freudian unconscious "directly into Lewinian language" by saying that "there are sub-systems within the region of the self which are not in communication with each other" (**5**, p. 225) is to slip over the specific nature of unconscious processes in Freudian theory. Furthermore, if this is meant as an equation, it fails to do justice to the very important communications which do exist between conscious and unconscious systems. The point may be illustrated by the dream, which draws upon the person's waking experiences (e.g., experiences of the "dream day") and which is recalled by the waking individual.

Another instance of a premature reconciliation of Gestalt psychology and psychoanalysis may be taken from the work of Witkin *et al.* (**18**).[6] These authors point out:

> Although psychoanalytic theory, in its conception of primary and secondary processes, recognizes the relation between intellectual functioning and personality, it has not really been concerned with the nature of secondary processes. . . . Gestalt psychology, in contrast, has offered a well-developed theory of cognition, in which the role of the nature of reality in determining perceptual and thought processes has been emphasized. But . . . Gestalt theory has on the other hand neglected the role of personal factors in perception. By showing that a perceptual act cannot be understood without reference to *both* personal factors *and* the nature of reality, studies such as ours help to bridge the gap between Gestalt and psychoanalytic theory, and provide a basis for bringing together the main aspects of both into a single comprehensive theory of human psychological functioning (**18**, p. 481).

If, as I believe, the authors are correct in saying that "there is still lacking in psychoanalytic theory any specific account of cognition or of the nature of secondary process" (**18**, p. 481), and that Gestalt psychology has neglected personal factors in perception—these are precisely the reasons why studies such as theirs do *not* help to bridge the gap between the two theories. The gap can be bridged only by a true reconciliation of existing differences; and since these two approaches have been concerned, as the authors point out, with such different areas of psychology, it is difficult even to know where the essential differences lie. A systematic analysis of the assumptions of both psychologies, one concerned with implicit as well as explicit assumptions, would undoubtedly reveal both important differences and surprising compatibilities of the theories. But in the absence of such an analysis we cannot, without glossing over real differences, say that a particular finding helps to bring the two theories together. To do so, it would need to be shown (*a*) that the results demand a theory which reconciles actual differences between the two approaches; or (*b*) if the two psychologies are in agreement in the area in question (which would require demonstration), that the findings can be handled in terms which are compatible with both.

To say, in other words, that both personal factors and reality factors determine a perceptual effect is to pose a problem. It is to point out that we need a human psychology which will include both kinds of factors; but it is not to say that such a psychology will be compatible with Gestalt psychology or with psychoanalysis or both. To the present writer it seems more likely that a finding which cannot be handled adequately within the framework of either of these existing systems will demand, not a reconciliation of the two admittedly incomplete theories, but rather a

[6] These remarks in no way detract from the excellence or the significance of these authors' study.

new theory. This point will be discussed below.

A final example, which will show again how eclectic reconciliation of differences may be achieved at the expense of a specific theory in the area of the controversy, may be taken from Welch (**17**). This author has offered some fundamental propositions which he believed should be acceptable both to Gestalt psychology and to contemporary behaviorism. For example, "perceiving is the result of a stimulus compound producing effects upon the sense organs which establish brain traces similar to or in otherwise related to brain traces formerly established" (**17**, p. 181). Surely everyone will agree that present percepts are related to traces of past ones; where dispute exists it concerns the specific effects of past experience on perception. Welch has succeeded in reconciling the differences by omitting the specific area of controversy. Again, we are told:

> In interacting with its environment, the organism changes in many ways. . . . [Among other changes] it may learn. Learning is the effect of a stimulus compound or stimulus compounds upon the nervous system of the organism and the responses which these evoke, that makes possible the establishment of new responses, as a result of such experiences (**17**, p. 187).

Many psychologists may, indeed, accept this as a rough definition of learning. But when one leaves this level of generality and raises the question of how this process is to be envisaged, this happy harmony disappears. Hilgard points out: "There are no laws of learning which can be taught with confidence" (**12**, p. 457). Likewise no one questions the fact that "behavior of any type is the result of the interaction of the organism and its environment" (**17**, p. 176). But what is the nature of this interaction? What are the roles of organism and environment? Here are questions on which different writers have taken divergent positions. (Cf. **11.**) While no one will disagree that memorizing and generalizing (**17**, pp. 181, 182) occur, and while Welch's definitions might provoke little controversy as rough identifications of the phenomena in question, different theories exist about the nature of these processes.

It will be clear from the above discussion that the existence of facts which all psychologists accept is irrelevant to the problem of eclecticism. Likewise the circumstance that some developments in psychology have called attention to facts ignored by others has no bearing on the issue. The important questions are: How are these facts understood? and What is their place in the overall theoretical system? Even where agreement exists as to the facts, differences are current with respect to these questions.

In all the examples considered here, it would appear that differences have been reconciled and controversy eliminated at the price of obscuring the issues with which research is concerned in contemporary psychology.

Boring, years ago, pointed out the productive role of controversy in scientific research (**2**). Not only does the eclectic lose prematurely the advantages of controversy, he may to some extent give up the advantages of theory as well. The above discussion contains the suggestion that the eclectic at times renounces specific theory in the area of a controversy in order to reconcile differences. This statement will be qualified below. But now attention must be drawn to a consequence of the intimate relation between fact and theory.

There is a certain amount of fact that can be discovered in the absence of any theory. For example, time errors forced themselves to the attention of psychologists who were concerned with quite different problems. For the most part,

however, problems for investigation arise out of the theories one holds. New facts are discovered in the course of research designed to test one's hypotheses. To the extent, therefore, that the eclectic gives up specific theory in the area of a controversy, he is handicapped in the discovery of new facts.

Closer examination will, however, often show implicit theories which may contradict the eclectic's avowed intention by placing him in a position on one side or the other of the (now only implicit) controversy. Woodworth, for example, deals with transfer in terms of *carrying over* something from one learning situation to another rather than in terms of *application* of what has been learned to the training tasks as well as to the new ones. That is, transfer is seen as occurring because knowledge acquired in the original training is carried over to the new activity; the new tasks, to the extent to which they are similar to the learned ones, are considered already partially learned. The alternative is ignored that what is learned is not tasks but principles or other whole properties; thus the training activity may merely provide examples of the use of the principle which can be applied equally to the new situation. Woodworth's theory is thus close to a theory of identical elements in this respect, opposing one derived from the study of learning by understanding (**13**, Chap. 5) and, indeed, unable to deal with many cases of such learning.[7] This consequence is particularly impressive since, as will be illustrated immediately below, Woodworth is by no means opposed to learning by understanding; it is another instance of the confusions which eclecticism breeds.

In another place Woodworth calls attention to the following controversy:

> Among present-day theories of learning those which emphasize reenforcement or the law of effect minimize the perceptual factor, often stigmatizing it as "mentalistic" and impossible to conceive in physical terms, while those which emphasize perceptual learning are apt to deny any direct importance to the factor of reenforcement (**20**, p. 119).

In attempting to show that there is "no obvious incompatibility" of these two factors, he makes (explicit) assumptions about learning as a cognitive process[8] which would be likely to be unacceptable to many S-R theorists, and (both implicit and explicit) empiristic assumptions about perception[9] which many

[7] If a theory of common factors were correct, there should never be more than 100 per cent transfer, since two activities cannot have more than 100 per cent of their factors in common. Yet, as Katona has shown (**13**), cases exist in which performance on the test activity is superior to that on the training task.

It is of interest to note also that Woodworth's theory, while it succeeds in reconciling the differences, appears to lump together cases that do not belong together. There is evidence that transfer of specific data is different, in process as well as in the magnitude of the effect, from the application of principles derived from one set of data to new material. (Cf. **13**.)

[8] For example: "As to connections, several may be established before the conditioning is complete, but the primary one connects the conditioned stimulus with the meaningful character it acquires as the first event in a regular sequence" (**20**, pp. 121–122). Also "In experiments that offer alternatives and demand a choice, what has to be learned is a distinction between stimulus-objects and not between motor responses. . . . What has to be learned is the difference between the two alleys" (**20**, p. 122).

[9] "When a new percept is in the making—when an obscure stimulus-complex is being deciphered, or when the meaning of a cue or sign is being discovered—an elementary two-phase process is observable. It is a trial-and-check, trial-and-check process. The trial phase is a tentative reading of the sign, a tentative decipherment of the puzzle, a tentative characterization of the object; and the check phase is an acceptance or rejection, a positive or negative reenforcement of the tentative perception" (**20**, p. 124).

Among the implicit assumptions seems to be the view that organization is not primary in

cognitive theorists might find equally unacceptable (assumptions, incidentally, which are not necessarily consistent with those about the learning process).[10]

Several questions suggest themselves with respect to the theory implicit in eclectic solutions. (*a*) A question worth examining is whether there is a tendency for such implicit theory to be too heavily weighted in the direction of traditional theory. As the above examples show, this need not always be the case; but it seems plausible to think that when theory is not explicit, and thus not examined, it draws upon doctrines prevailing both in psychology and in the culture in general rather than upon the newer and less widely accepted theoretical currents. In a similar connection Köhler has pointed to a certain conservatism in eclecticism:

> . . . it has been said with approval that psychology now tends to be eclectic. Again, we have been told that in psychology we had better stay in the middle of the road. I cannot agree with these prescriptions because, if they were followed, psychologists would have to look first of all backward. In an eclectic attitude, they would be too much concerned with ideas which are already available; and, in attempting to find the middle of the road in psychology, they would have to give too much attention to the tracks along which others have moved before them. Such attitudes could perhaps be recommended if, in research, security were an important issue. Actually there is no place for it in this field. In research, we have to look forward, and to take risks (**15**, p. 136).

(*b*) Another question which arises in connection with the theory underlying eclectic solutions is the following: since such theory is often implicit, and thus unexpressed and unexamined, is it adequate to lead to the discovery of new facts? For example, since the idea of "carrying over" (i.e., as opposed to that of "application") is only implicit in Woodworth's theory of identical components, it seems unlikely that it would be subjected to test. Or again, the implicit elementarism in Welch's statement about learning (cf. Footnote 10) is unlikely to be tested since the author's main focus is on other aspects of the statement.

Also worth looking into in connection with the theory implicit in eclectic solutions are the questions of its adequacy for ordering the facts and its susceptibility to proof or disproof. For example, to say that "something" is transferred is too unspecific a statement of the conditions of transfer to test empirically. Any finding of transfer seems to confirm it, and there is no result which could disprove it. Again, it has been suggested above that the theory implicit in a given eclecticism is not always internally consistent. This is a question which deserves examination in connection with particular eclectic psychologies.

We may summarize the discussion so far by saying that eclectics have to a large extent succeeded in resolving conflicts in psychology by ignoring differences and obscuring the issues. Some reasons for dissatisfaction with such solutions have been indicated.[11] Is there no

perception, nor prior to the effects of learning; as well as the idea that there is no fruitful distinction to be made between perception and interpretation.

[10] As a final illustration, Welch states: "This distinction between elementary and higher forms of learning involves the distinction between a situation where the new elements are simple in nature, or simple in character and are simply integrated, and a situation where the new elements are complex and integrated in a complex manner" (**17**, p. 188). This statement implies an elementaristic view of the learning process—learning being envisaged as the integration of elements—which would be far from acceptable to all the psychologists Welch is trying to reconcile.

[11] It is of interest to note that eclecticism seems to have presented similar problems in other fields of knowledge in their comparative

alternative? It seems to the present writer that reconciliations can be reached in psychology only by focusing on the existing differences, examining them, and carrying on research to settle issues. If this is eclecticism, it is eclecticism after the fact rather than the prevailing eclecticism before the fact. And it is clear that it will not be a matter of reconciling existing theories. Since competing theories on any particular issue in psychology today—or competing psychological systems—each tend to be plausible and to be supported by evidence, it is unlikely that any one will win a clear victory over the others. Yet none can offer a fully satisfactory explanation—or else the controversy would not exist. Controversies do not exist in science with regard to processes which are fully understood. Thus the task seems to be one of arriving at new, more comprehensive theories of the processes in question.

youth. I quote an observation on the medical science of a century ago: ". . . And as the rules derived from fundamental truths seemed to come into unsolvable contradiction with the experiences and the sanctioned standards of practice, there sprang up under the name Eclectic the representatives of sober elucidation, of the *juste milieu,* of the medium of the extremes. The breach between theory and practice, which they feared, was avoided or postponed if theory gave up the pretension to penetrate into particulars and if practice agreed that, because of its youthful immaturity, it should be excluded from counsel, and progress in silence and in hope. The conflict was settled and peace was achieved, not by the reconciliation of the parties, but by separating them. The so-called impartial examination of the facts should lead only to a middle road between them. [The eclectics] thought they had principles and avoided their application; they proclaimed themselves free and in practice clung to the consequences of old dogmas. They practiced tolerance not because they included the truth of each dogma, but because a chasm existed between theory and life, beyond which theory didn't matter" (**9**, p. 9).

An example should make this clear. It seems safe to say that theories of forgetting arising out of experimental psychology have found no adequate place for the facts of repression. Nor have the psychoanalysts succeeded (or tried) to bring these facts into relation with a general theory of memory and forgetting. Can the two kinds of theories be brought together? It seems to me that the most fruitful starting point is not the attempt to reconcile existing theories. Actually, useful theories of repression do not exist. (Cf. **6** for a similar point, more generally stated.) It is hardly sufficient to say:

> Repression proceeds from the ego, which, possibly at the command of the superego, does not wish to be a party to an instinct cathexis originating in the id. Through repression the ego accomplishes the exclusion from consciousness of the idea which was the carrier of the unwelcome impulse (**7**, p. 19).

This statement contains no hypothesis about the processes involved, about how repression can possibly be brought about. Thus there seems to be no point to attempt to reconcile the theories of experimental psychology and of psychoanalysis on repression; neither has an effective theory in this area. What we need is to look into the processes themselves, in the light of what we know about forgetting in general (cf. [**10**]). Can affective processes act, for example, to produce a failure of the Höffding function—i.e., that selective interaction between present process and memory trace which is the basis of recognition and the first step in the process of recall? (Cf. **14**, pp. 126 ff.) Under what conditions can emotional and motivational processes introduce interferences? Answers to such questions might lead not only to a hypothesis about the nature of repression, but might also introduce considerable modification into our

present theories of the nature of forgetting in general.

The eclectics are, of course, right in maintaining that where a genuine controversy exists in psychology, and where evidence seems to support both sides, there is likely to be some truth to both positions. But they solve their problem too soon. Existing theories cannot be made more comprehensive by adding divergent ones together. They can be broadened to include all the relevant evidence only by looking more deeply into the phenomena with which they are concerned; and this means arriving at new theories.

At this point the parallel between productive solutions of theoretical problems and of personal problems becomes striking. In connection with the reconciliation of opposites within the personality, C. G. Jung points out that conflicts are never resolved on their own level. They are outgrown. Only on a higher level can you see both sides.

Summary

Examples have been presented to show that eclectics tend to resolve conflicts in psychology by glossing over real differences and obscuring the issues. Such solutions achieve harmony at the price of specific theory in the area of the controversy, and thus sacrifice fruitfulness in the discovery of new fact. Closer examination often reveals implicit theories underlying such solutions, but unexpressed and unexamined theory can hardly be expected to equal explicit hypotheses either in fruitfulness or in adequacy in dealing with known facts.

It is here suggested that differences need to be resolved in psychology not by denying them and attempting to combine existing theories, but by focusing on the differences and using them to get a better view of the relevant phenomena. We will achieve more comprehensive theories not by combining existing ones but by understanding better the processes in question.

REFERENCES

1. Abt, L. E. A theory of projective psychology. In L. E. Abt & L. Bellak (Eds.), *Projective psychology*. New York: Knopf, 1950.
2. Boring, E. G. The psychology of controversy. *Psychol. Rev.*, 1929, **36**, 97–121.
3. Boring, E. G. *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts, 1950.
4. Boring, E. G. The nature of psychology. In E. G. Boring, H. S. Langfeld, & H. P. Weld (Eds.), *Foundations of psychology*. New York: Wiley, 1948.
5. Bronfenbrenner, U. Toward an integrated theory of personality. In R. R. Blake & G. V. Ramsey (Eds.), *Perception, an approach to personality*. New York: Ronald Press, 1951.
6. Bruner, J. S. Freud and the image of man. *Amer. Psychologist*, 1956, **11**, 463–466.
7. Freud, S. *The problem of anxiety*. New York: Norton, 1936.
8. Freud, S. *Beyond the pleasure principle*. London: Hogarth Press and the Institute of Psycho-Analysis, 1950.
9. Henle, J. *Handbuch der rationellen Pathologie*. Erster Band. (2nd ed.). Braunschweig: F. Vieweg u. Sohn, 1846.
10. Henle, Mary. Some effects of motivational processes on cognition. *Psychol. Rev.*, 1955, **62**, 423–432.
11. Henle, Mary. On field forces. *J. Psychol.*, 1957, **43**, 239–249.
12. Hilgard, E. R. *Theories of learning*. (2nd ed.) New York: Appleton-Century-Crofts, 1956.
13. Katona, G. *Organizing and memorizing*. New York: Columbia Univer. Press, 1940.
14. Köhler, W. *Dynamics in psychology*. New York: Liveright, 1940.
15. Köhler, W. The scientists and their new environment. In W. R. Crawford (Ed.), *The cultural migration*. Philadelphia: Univer. of Pennsylvania Press, 1953.
16. Lewin, K. *Principles of topological psychology*. New York: McGraw-Hill, 1936.
17. Welch, L. An integration of some fundamental principles of modern behavior-

ism and Gestalt psychology. *J. gen. Psychol.*, 1948, **39**, 175–190.

18. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception.* New York: Harper, 1954.
19. WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.
20. WOODWORTH, R. S. Reenforcement of perception. *Amer. J. Psychol.*, 1947, **60**, 119–124.
21. WOODWORTH, R. S. *Contemporary schools of psychology.* (Rev. ed.) New York: Ronald Press, 1948.
22. WOODWORTH, R. S., & MARQUIS, D. G. *Psychology* (5th ed.). New York: Holt, 1947.
23. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology.* (Rev. ed.) New York: Holt, 1954.

(Received January 9, 1957)

Psychological Review
Vol. 64, No. 5, 1957

VISUAL PERCEPTION: AN EVENT OVER TIME

GUDMUND SMITH [1]

Research Center for Mental Health, New York University

Most psychologists would certainly agree that perception—or more broadly, cognition—like other aspects of behavior, is not an instantaneous datum but, instead, that it should be considered an event over time. There are only a few of them, however, who have consistently entertained this self-evident notion when designing their experiments and their explanatory models. Until recently only the Leipzig psychologists, in their study of the *Aktualgenese* of percepts, had tried to make the microgenesis of perception the center of their experimental and theoretical approach. Their admitted failure to call attention to the genetic point of view must in part be attributed to their method of developing concepts which were mainly of a descriptive order (for further criticism, also methodological, see 8), and in part to the inability of current schools of psychology to include the Leipzig findings in their explanatory systems.

This paper is not intended as an attempt to rehabilitate the Leipzig school, however justified that may be, but an attempt to reconsider the genetic approach to problems of visual perception. In view of recent contributions within the fields of personality and perception, such an approach may seem more rewarding than it did thirty years ago; Hebb, for instance, in his physiological theory, takes into account the "phase sequence" as a prerequisite for the organization of a percept (5). The present author and his collaborators have tried to demonstrate, more specifically, that "serial" interpretations of perceptual reports continuously repeated in the same experimental situation (a situation which should be so new or so complicated as not to allow an adequate report after only one trial) often tended to reveal more about an individual's adaptive mechanisms than interpretations only of the final, stabilized reports or of the usual summary scores (10, 13). Our basic consideration thus implies that we look upon perception (cognition) as a process of organization, emphasizing that it needs time to be prepared, evolved, and established. This consideration has been developed in earlier papers (e.g., 11) and partly in cooperation with Kragh (8), although some points where the present author differs from him will be mentioned later (cf. 12).

Basic Assumptions

The process of visual perception should not be confused with what is generally called percepts, which are products of the process fitted into a frame of reference of outside reality. It would be even more correct to say that percepts normally are the products of late phases in a process of organization, the preparatory phases of which are therefore valid objects of study. Some important characteristics of these early or preparatory phases are implied already in the basic consideration. First of all, we have to assume that these stages do not generally produce percepts; only rarely do we become aware of our own perceptual processes. This assumption is of extreme importance for any genetic theory of perception, since it indicates that we do not have to conceptualize preparatory stages in terms of finished products, e.g., as faint copies of established percepts or of stimuli, or as reflections of physiologi-

[1] Now at the University of Lund, Sweden.

cal or other processes hypothesized on the basis of end-product behavior (11). It is more natural to assume that preparatory stages often differ from end stages just because they have not been able to produce percepts. Such premature products of early stages would probably be suppressed, at any rate, in favor of something more suitable for the frame of reference of outside reality.

When the perceptual process is inhibited, as in a tachistoscopic experiment where exposure times are too brief for a correct perception of the stimulus, early stages of the perceptual process can be forced to produce a percept. But we know that these percepts are often diffuse and equivocal. Another characteristic of preparatory stages might therefore be that they embrace numerous "possibilities" for further development. Let us, to simplify matters, call it an unspecific attention toward something that is going to emerge (e.g., to be seen), a mobilization of a system for many possible activities. As the ideal perceptual process continues in a situation with prolonged stimulation, however, more and more of these possibilities fall into the background in favor of the one possibility that coincides with the stimulus (or, with a normative conception of the stimulus), in favor of the unequivocal reality of the individual (or his culture, or his language). Only fairly unequivocal stages can produce percepts, at least in so-called normal human beings. We should then state, more generally, that *one* possibility has to be singled out for the perceptual process to produce a percept—and, of course, not always the possibility that is correct from the point of view of stimulus. We know well by now that even "normal" people can perceive incorrectly.

Now we can be a little more specific about some additional characteristics of the preliminary stages. It seems reasonable to assume that it is not impossible for them to harbor completely contradictory possibilities of development, i.e., contradictory from the point of view of outside reality (stimulus). When reversible figures are presented near or below the threshold of awareness, many individuals seem to recall both "sides" of them at the same time; [2] or they report the correct percept as well as various incorrect ones, the full-fledged Gestalt along with fragments of it, etc. Although we conceive preliminary phases as encompassing contradictory possibilities "side by side," we should not think of them as a number of superimposed percepts. As stated above, the preparatory stages should not be conceptualized in terms of established ones. It would probably be more fruitful to try to relate preparatory phases to final ones in the same formal way as primary processes are related to secondary ones in psychoanalysis; i.e., only the final phases have to adhere to the logic and organization of outside reality. A later stage, however, cannot be fully understood with reference only to this outside reality, to the stimulus. The main emphasis, from the genetic point of view, is upon the sequence over time, the interpretation of late stages in the light of earlier ones. Even if the later stage is different in scope and organization from previous stages and, emerging on top of the sequence, is in control of them, the roots of this stage are still to be found in the preparatory part of the perceptual process.

THE *AKTUALGENESE* APPROACH

The Leipzig school made an attempt to capture the perceptual process by presenting the same stimulus, first at very brief exposures, then at gradually prolonged ones. Alternative methods were also employed, e.g., when the intensity

[2] Charles Fisher: Lecture at New York University, 1956.

of the stimulus was increased from one trial to another. It was assumed that the shorter the presentation, or the lower the intensity of the stimulus, the earlier the stage at which the process was forced to produce a percept. Kragh (8) has refined the tachistoscopic method considerably: he used new types of stimulus material, e.g., TAT-like pictures, and tried to define the "fractionized" stimulus more rigorously than before. But it is even more important that he developed his serial or microgenetic method in close relation to a conceptual model of perception-personality. The two concepts supporting his model are ***construction*** and ***reconstruction***. Construction refers to the evolving process of organization, the ideal end product of which is the correct percept. But our knowledge about such a process is always based upon a series of reports of gradually prolonged stimulation. In order to describe the temporal organization of a percept we thus use the method of reconstructing stages before the final stage; i.e., various phases of the process of construction have been enticed prematurely to produce percepts fitted to the conceptual level of the end product. Nevertheless, a series of reports of a stimulus exposed for gradually extended times should, according to Kragh, reflect the microscopically short genesis of a percept even if the filtering process of reconstruction deprives the early stages of some of their qualities. Since such a series of preliminary reports implies a continuous progress toward the final percept, however, qualities lacking in early stages may frequently be inferred from later stages.

It is not our intention to question here the assumption that these perceptual reports belong to the *same* process of organization. But we are less optimistic than Kragh as to the possibility that a report in his experiment can be regarded as a veridical representation of the early stage at which the perceptual process was supposed to have stopped. Kragh's concept of reconstruction pays attention to the fact that, in *Aktualgenese* experiments, early stages have to be communicated on an end-stage level, but it does not refer explicitly to the equally important fact that, since percepts have to fit into reality, a number of other perceptual possibilities must have been disguised or omitted in order to facilitate the event of a percept. Thus, in *Aktualgenese* experiments, a series of early reports are likely to reflect the early part of the perceptual process very insufficiently, or at least in the same complicated manner as dream condensations reflect primary as well as secondary processes. But this does not preclude the assumption that they represent very powerful possibilities of development which may manifest themselves in one experiment after the other, only slightly disguised. One possibility of this kind would naturally be the correct percept, as was demonstrated by the compulsive group in Kragh's experiments (8); these people tended to proceed directly from blank stages to correct ones, i.e., to avoid or isolate all incorrect possibilities. In other subjects it was evident that many of the alternative perceptual possibilities were related to important personal experiences far back in life.

Another important assumption is that the perceptual process developed over an intermittent series of short presentations is equivalent to the process accompanying one long presentation of the same stimulus. Kragh maintains that there is an approximate similarity between everyday perception and perception in *Aktualgenese* experiments, partly because our fixation even of immovable objects is likely to be intermittent. But these intermittent fixations in normal perception are not accompanied by conscious

perceptual products differing in degree of clarity and correctness; i.e., while tachistoscopic experiments tend to *fix* the perceptual process at various stages of organization, the normal process is probably more continuous. It seems only reasonable to assume that the course of perceptual organization will change more or less drastically when one of its erroneous alternatives appears as a conscious percept. The *Aktualgenese* method also tends to hamper the feed-back mechanics which probably play an important role in normal perception (1). These and previous considerations regarding the influence exerted on the perceptual process by the *Aktualgenese* method may be important for the assessment of conclusions based on *Aktualgenese* data. But they give us no reason to abandon the working hypothesis that perception is a process of organization extending over microtime.

ELABORATION OF SOME BASIC ASSUMPTIONS

One of the basic assumptions for this model of perception is that the "germ" of a correct percept is embodied already in early stages of the perceptual process. Naturally, this would be true only in cases where, from the stimulus point of view, a correct perception was possible, i.e., where the retina was able to register a sufficiently large part of the object, etc. If this was not the case, the perceptual process would have to be reinforced by way of a new fixation. Any repeated or prolonged fixation would certainly strengthen the correct possibility. Let us assume, at the same time, that early stages are more prone to produce ideas, dreams, etc., than percepts—an assumption in complete accordance with previous assumptions about these stages. It is hardly surprising, then, that in dreams which their patients reported on the morning after the experiment, Poetzl (9) and Fisher (4) found exact copies of those parts of the stimulus that had not been consciously perceived together with fragments of them. We are convinced, furthermore, that many more "distorted" products of the same process would be found in these dreams if there was any acceptable method to account for them. The more we reduce the intensity or duration of stimulus the easier it will be for "incorrect" products to appear instead of the correct one. And when the stimulus itself is equivocal, as in many projective tests, there is no longer any "correct" possibility.

The exact nature of the perceptual process hypothesized here may still seem a little mysterious. The primary source of knowledge is perceptual reports, either reports of end stages obtained under optimal stimulus conditions or of such quasi end stages as imply that the perceptual process has been prematurely interrupted. An alternative to the latter methods will be described later in this paper. But these sources are still insufficient to account for the perceptual process. If our main postulates are correct we can learn much about it, especially its early phases, by studying such products as dreams, images, etc., i.e., products that cannot be accepted as outside reality here and now. In this way we bring the study of perception into the broader field of personality. We may even say that the process of perceptual organization is a reflection of personality in the same way as its end products are reflections of the reality built up around the individual. When we thus postulate that we can learn much about basic processes in the individual by reconstructing the construction process of perception, we have not implied anything mysterious or even radically different from current conceptions about perception-personality.

PERCEPTION-PERSONALITY

The process of perceptual organization may be described essentially in formal terms or in terms mainly of content. It is expected to show characteristics typical of the individual, so-called consistencies or structures, and also characteristics closely related to the current situation. Some structures would be more typical of the early phases of the process than of later ones because the objectivized conception of reality is similar in most individuals; this possibility has been realized in much present-day perception-personality research where the intensity or clarity of the stimulus is cut back. But this would not be enough to account for the structures reflected in the perceptual process. Structures also refer to the longitudinal aspects of the process, the way in which certain end stages are reached and others warded off, to all those aspects of control of the organization process which are termed cognitive attitudes or system principles (**2**, **6**). For instance, while the perceptual processes of compulsive people seem to be controlled by an all-or-nothing principle, by avoidance of all stages between a blank report of stimulus and a correct one, those of other people tend to develop more continuously over a series of more or less incorrect (subjective) reports (**8**, **15**). However, we do not have to explain these and other characteristics of the process by referring to constructs outside the process, or method, itself. The explanatory constructs are given in the process of construction and the operation of reconstruction by which this process is unveiled.

It would be even more correct to say that the explanatory constructs of any temporarily limited process of organization are ultimately given in that all-embracing continuous process of construction the qualities of which we find in the life history of the individual. A perceptual process is naturally to be regarded as a continuation of this chain of adaptive events. Then, in its various stages the perceptual process must also be marked by primary or secondary characteristics of personality organization; the incorrect "possibilities," for instance, are not accidental (how could they be?), but often reflect those experiences out of which the current individual was shaped (**8**). Thus perception has the same roots as other forms of behavior, although its end products manifest themselves in the specific frame of reference of outside reality. We may go even further and assert that percepts and some aspects of simultaneous overt behavior are two sides to the same event, an assumption implied in our common acceptance of verbal reports as indicative of percepts. But while numerous unsuitable products of a process of organization cannot manifest themselves as percepts, they may appear in other forms of behavior which do not have to be so closely fitted into a normalized framework. And these reflections of preparatory stages, as said above, are most obvious in such products as were discarded in the perceptual process as unsuitable representatives of outside reality, viz., our own dreams, images, hopes, phantasies and the like (see also below concerning the measurement of nonverbal forms of behavior). Only by studying all behavior manifestations can we learn the full story of perception.

PERCEPTION BEYOND AWARENESS

The problem of perception beyond awareness must be considered crucial for the approach outlined here, which implies that an exhaustive study of the problems of perception cannot be restricted to the characteristics of conscious end products. In collaboration with a number of other psychologists, the present author has tried a new tech-

nique for investigating preconscious aspects of the perceptual process. This technique has been reported in detail elsewhere (**7**, **14**, **15**, **16**). The following procedure may be regarded as typical (cf. also **3**, **17**). Two stimuli are exposed in rapid succession in a tachistoscope. The exposure time for the second of these stimuli (*B*) is so long as to admit of a correct, stabilized perception of it, while the first stimulus (*A*) is presented too briefly to be conceived as an independent entity by the subject. Since *A* and *B* appear on the same part of the stimulus field, one can say that the perception of *A* is extinguished by the perception of *B*. But in a number of subjects the subliminal perception of *A* affects the perception of *B* in various ways; the two percepts seem to coalesce into a changed perception of *B*. This is to say, we study a perceptual process (*A*) beyond awareness by reading off its effects on an adjacent, conscious percept.

In some of our studies, the phase of the *A* process influencing the perception of *B* seemed to have been comparatively late, at least so late that the stimulus "possibility" asserted itself beyond other possibilities. This was evident in an experiment where a fan-shaped line pattern (Stimulus *A*) influenced a square (Stimulus *B*) as if the former subliminal stimulus had been correctly perceived; i.e., the square developed into the same rhomboid as in the well-known illusion with *B* on the lined background of *A*. But although the *A* process thus changed the reports of *B* in a direction predictable from the correct *A*, it was still too weak and diffuse to assert itself as a full-fledged percept beside the *B* percept. If it should be represented at all it had to be represented within the context given by the latter percept. However, while this experiment was designed to prove only that the perception of *B* may be affected by a perceptual process beyond awareness, later experiments indicated that the relations between the preconscious process and associated conscious reports differed interindividually (**15**, **16**). Here stimuli were more complex and were made up of meaningful line drawings. In other words, the characteristics of early stages of the perceptual process, as implied in reactions to *B*, did not reflect only characteristics of the subliminal stimulus but also reflected basic cognitive attitudes, e.g., the defence mechanisms of projection and isolation. Thus in a paranoid group the subliminal influence of *A* upon the perceived reality context (*B*) was often drastic, and varied from one individual to another, whereas in a compulsive group the *A* process seemed to be isolated from *B* until *A* emerged as a correct, independent entity (**15**).

It seems rather self-evident that a vague and equivocal conscious percept would be more easily affected by a subliminal process than would a clear and well-structured percept. In an experiment with a complex *B*, a face, we found greater effects in the beginning of a series of reports than later, when the impression of the face became more stabilized; only when the exposure time for *A* approached threshold values did its influence on the perception of *B* increase again (**16**). It was interesting to note that the effect of subliminal stimulation (the words HAPPY and ANGRY) was not necessarily directly related to the conventional meaning of these words; i.e., ANGRY did not produce only angry expressions in the face but more often also tense, anxious, serious, pensive, and similar expressions. This may partly confirm our assumption that the early stages of a perceptual process include more varied "possibilities" than those related to conscious, stabilized cognition, though we also have to consider that interaction between the subliminal process and con-

scious thoughts about the face probably contributed to the modification of meaning. If Stimulus *A* is very novel, unusual, or controversial, there is reason to believe that, in many individuals, other possibilities than the correct one will influence the *B* percept, or, in other words, that the difference between sub- and supraliminal conditions will be considerable (**7**). The same will be true of differences between individuals in the subliminal condition.

In most of these experiments on perception beyond awareness, the actual descriptions constituted the basic core of information. But if subjects could become sufficiently relaxed, images and by-remarks accompanying the reports might be even more important, as demonstrated by Kragh (**8**), who used a free-association technique in his studies. *A* tendencies which cannot assert themselves within *B*, and which cannot be represented within this frame of outside reality, may very well appear in images, associations, denials—or, for that matter, they may influence other, nonverbal forms of behavior. The study of these behavioral aspects should become more crucial the more resistant to change are perceptions of *B* stimuli. In one of the studies just mentioned, for instance, differences in latency of reactions to a *B* stimulus presented in alternating combination with the subliminal stimuli ANGRY and HAPPY seemed to *increase* as differences in verbal description of *B* in these two pairings *decreased*. In other words, where the two different *A* stimuli did not produce differences in a subject's descriptions of *B*, a face, they produced differences in the time taken before he started to react verbally (**16**). Dreams reported the morning after the experiment have also proved important to account for the structures and meanings in the preconscious *A* process. However, only by extending these studies to still more general aspects of personality organization can we fully assess the implications of our findings. A preliminary attempt was made in the paper just referred to.

The experiments described here naturally represent only a few variations of the basic method which may be applied in many forms to a variety of perceptual phenomena. Generally, I think that in these experiments, where the *A* process does not have to manifest itself as independent conscious percepts, it will be possible to explore very important aspects of the perceptual process that cannot be reached in experiments of the *Aktualgenese* type. One of the specific problems raised by the new method concerns the interrelations between a perceptual process, the products of which hold the center of awareness, and more marginal processes continually evolving at the same time—interrelations of immediate relevance for the study of everyday perception and, as has been illustrated here, of basic mechanisms of control and adaptation. But most important of all, both types of experiments imply a challenge to the traditional assumption that perception is an instantaneous event the mechanisms of which are reflected only in conscious, unequivocal products.

SUMMARY

In this paper perception is considered as a microscopically short process of organization, the prestages of which are therefore important objects of study. Some theoretical and empirical approaches to this genetic analysis of perception (and personality) are discussed together with the topic of subliminal perception, which is of particular importance in this connection.

REFERENCES

1. ALLPORT, F. H. *Theories of perception and the concept of structure.* New York: Wiley, 1955.

2. ANGYAL, A. *Foundations for a science of personality.* New York: Commonwealth Fund, 1941.
3. CHEATAM, P. G. Visual perceptual latency as a function of stimulus brightness and contour shape. *J. exp. Psychol.*, 1952, **43**, 369–380.
4. FISHER, C. Dreams and perception. The role of preconscious and primary modes of perception in dream formation. *J. Amer. Psychoanal. Ass.*, 1954, **2**, 389–445.
5. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.
6. KLEIN, G. S. Perception, motives, and personality. A clinical perspective. In J. L. McCary (Ed.), *Approaches to personality.* New York: Logos Press, 1956. Pp. 121–199.
7. KLEIN, G. S., SPENCE, D. P., HOLT, R. R., & GOUREVITCH, SUSANNAH. Cognition without awareness: subliminal and supraliminal influences upon conscious thought. *J. abnorm. soc. Psychol.*, in press.
8. KRAGH, U. *The actual-genetic model of perception-personality: an experimental study with non-clinical and clinical groups.* Lund, Sweden: Gleerup, 1955.
9. POETZL, O. Experimentell erregte Traumbilder in ihren Beziehungen zum indirekten Sehen. *Z. Neurol. Psychiatr.*, 1917, **37**, 278–349.
10. SMITH, G. Twin differences with reference to the Müller-Lyer illusion: a study in modal summaries and serial change. *Lunds Universitets Årsskrift. N.F. Avd. 1*, 1953, **50**, 1.
11. SMITH, G. The place of physiological constructs in a genetic explanatory system. *Psychol. Rev.*, 1954, **61**, 73–76.
12. SMITH, G. A new psychological model. *Theoria*, 1956, **12**, 61–69.
13. SMITH, G., & KLEIN, G. S. Cognitive controls in serial behavior patterns. *J. Pers.*, 1953, **22**, 188–213.
14. SMITH, G., & HENRIKSSON, MAJ. The effect on an established percept of a perceptual process beyond awareness. *Acta Psychol.*, 1955, **11**, 346–355.
15. SMITH, G., & HENRIKSSON, MAJ. Studies in the development of a percept within various contexts of perceived reality. *Acta Psychol.*, 1956, **12**, 263–281.
16. SMITH, G., SPENCE, D. P., & KLEIN, G. S. Subliminal effects of verbal stimuli. *J. Pers.*, in press.
17. WERNER, H. Studies on contour: I. Qualitative analyses. *Amer. J. Psychol.*, 1935, **47**, 40–46.

(Received January 3, 1957)

Psychological Review
Vol. 64, No. 5, 1957

RECENCY, FREQUENCY, AND PROBABILITY IN RESPONSE PREDICTION

JOHN E. OVERALL AND W. LYNN BROWN [1]

The University of Texas

It has been ten years since Voeks (13) aroused psychological interest by showing that contemporary learning theory can be used as a basis for individual response prediction. It still remains a point of controversy as to whether *S*s utilize all past experiences in the situation (frequency), or whether only the most recent experience determines the response (postremity). Voeks found that postremity predictions were the more accurate, even where frequency and postremity led to opposing predictions. Nevertheless, it was observed that postremity predictions were less accurate where they conflicted with frequency predictions than where the two did not conflict. This suggests that a theoretical model which bases predictions upon *both* frequency and recency might lead to even more effective response prediction.

Theoretical Models for Individual Response Prediction

Contemporary learning theories may be divided on the basis of adherence to a recency principle or adherence to a frequency principle. A brief survey of several contemporary theories will make this problem explicit.

Guthrie's Postremity Predictions

One group of theorists, whose leading exponent is Guthrie, sees recency as the important determiner of behavior. Guthrie says that stimuli and responses become associated in an all-or-none fashion by a single contiguous occurrence. There is no problem of strengthening; either the response is associated full strength with the stimulus or it is not. When the stimulus next occurs, the response made last in its presence will be repeated. Voeks (14) has formalized this in her principle of postremity. ". . . A stimulus which has accompanied or immediately preceded two or more incompatible responses is a conditioned stimulus for only the last response made while the stimulus was present" (6, p. 344). For the T-maze situation to be investigated in this paper, postremity would seem to lead to the prediction that the *S* will choose the same cul he chose last in the situation.

Postremity-Reward Predictions

Postremity yields a different type of prediction if the incentive value of reward is considered. Where the noncorrection method is employed and no reward follows a response, the very last movement made, after it is discovered that no reward is present, might be a "fractional" movement toward the opposite goal box. This would lead to the prediction that *S* will respond on the next trial by going where he went last, if he found reward, but that he will respond on the next trial by going to the opposite side, if no reward were encountered on the last trial.

Predictions Based Upon Frequency of Response

In sharp contrast to the Guthrian position are theories which say that all previous experiences in the situation enter into the response determination.

[1] The cost of this investigation was met in part by a grant from the Research Institute of The University of Texas, Austin, Texas.

Such theories neglect the principle of recency entirely.[2] Response predictions based upon frequency of response might be derived from Spence (9), who classifies maze learning as instrumental behavior. Since reward is not assumed to work backwards to strengthen associations made along the path to the reward, and since it is assumed to function through fractional anticipatory responses as an *incentive motivator* only, then it follows that increments of associative strength (*sHr*) accrue simply as a function of number of occurrences. The more frequently an instrumental response is executed, the more associative increments are added to the strength of the habit. The importance of recent events in the learning sequence is not emphasized. Thus, frequency of response provides another basis for predicting individual behavior.

Frequency of Reinforcement

If reward is assumed to reinforce instrumental acts along the path according to a goal gradient, then frequency of reward may be employed as another basis for predicting individual responses in a selective learning situation.

Algebraic Sum of Positive and Negative Cases

The writings of Hull (7) provide still another model for predicting individual responses, although he pleads that the concept of oscillation will prevent individual response prediction from ever achieving high accuracy. Hull classifies learning in a T maze as a kind of trial-and-error learning. Events at the goal box strengthen reactions made in arriving at the goal according to a diminishing function known as the "goal gradient." Each time reward follows the response, an increment is added to *sEr* associated with that response, and each time a response is not followed by reward an increment is added to *sIr* associated with the response. Hull states: ". . . We also tentatively assume, though without adequate evidence, that ΔsIr follows the same law with the same constants as ΔsEr" (7, p. 24). The effective reaction potential associated with one alternative in the T maze may be roughly estimated as $sEr = sEr + sIr$, where increments to *sEr* and *sIr* are scaled directly in terms of reinforced and nonreinforced outcomes. This procedure admittedly neglects "generalization" of both *sEr* and *sIr*, as well as the postulated progressive reduction in increment magnitudes as *sEr* and *sIr* approach their respective maximums. The defense for this procedure is that in a random 50–50 T-maze situation these factors should very largely cancel out, with generalization roughly equal for the two alternatives and with the reduction in increment magnitudes for *sEr* matched by a similar reduction for *sIr*. Consequently, we have still another frequency-type model for predicting individual responses—i.e., the algebraic sum of positive and negative instances associated with each response alternative.

Probability of Reward

Another theoretical model which yields individual response predictions is a cognitive-probability model. Tolman and Brunswik (12) conceived of the organism as relating *signs* to *means-objects*, which in turn are "causally" related to *desired goals*. However, because the causal texture of the environment is not univocal, the organism develops hypotheses based upon the probabilities of

[2] It should be noted that leading frequency theorists attempt to sidestep the problem of individual response prediction by pointing to lack of control and inadequacy in experimental techniques. It cannot be denied, however, that learning is a phenomenon which takes place within the organism; consequently, the final test of a theory of learning must be its ability to predict and make understandable the specific behavior of individuals.

means-objects resulting in *desired goals.* That route or behavior is chosen which has the highest probability of leading to reward. By considering the relative frequency of rewarded responses to total responses

$$\left(p = \frac{\Sigma R}{\Sigma R + \Sigma N}\right)$$

for each T-maze alternative, it is possible to calculate mathematically a probability estimate associated with each alternative. If this is done before each individual response, it may be predicted that *S* will choose the alternative with the greater probability of reward.

Recency-Weighted Probability Model

As outlined above, the probability model considers the recency principle no more than does a frequency model. However, it is tentatively suggested that the probability model might be revised in a manner compatible with the general theory of Tolman (11), especially to include capacities of the organism such as "retentivity," which Tolman emphasizes in his "capacity laws." This revised expectancy theory should provide the basis for more accurate response predictions which are based upon a weighting of *both frequency and recency.* It is recognized that the recency principle might be incorporated into contemporary frequency theory as easily as into the cognitive model developed here.

Three steps in the determination of a single selective learning response will be considered. First, the initial nature and magnitude of memory traces must be established. Past events in the learning sequence can affect present responses only through traces left by them within the nervous system of the organism. Secondly, traces within the memory system are not permanent; consequently their decay with time must be considered. Finally, selection of the response will be considered.

Creation of initial traces within the memory system. The probabilities of reward and nonreward upon which responses are based must be determined by past experiences of reward and nonreward in the situation. Since past experiences are themselves no longer present, it becomes important to understand the process whereby *traces* of past events are created within the memory system of the organism. Traces within the memory system are conceived to be positive or negative and to have greater or smaller magnitudes as a function of "expectancy."

It is proposed that the initial trace created within the memory system by a behavior outcome will be *positive* if that outcome is more "desirable" than the expected outcome, and will be *negative* if the outcome is less "desirable" than the expected outcome. An outcome which is exactly equal to the expected outcome will be neither positive nor negative. For example, Tinklepaugh (10) found that a lettuce leaf which would have had a positive valence under certain other conditions was rejected as negative by his chimps when they expected to find a banana. Crespi (2, 3) obtained different performance values as a function of expectancy in partial and in 100 per cent reinforcement situations, and Elliott (4) demonstrated shifts in maze performance of rats resulting from changes in the type of food reward.

It is proposed that the initial *magnitude* of a positive or negative memory trace is a function of expectancy. The farther a behavior outcome departs from the expected outcome, the greater the magnitude of the trace created by it within the memory system. Thus, if the expectancy of reward is low, a single rewarding outcome will create a trace of relatively large magnitude. In a situation of low-reward expectancy a single nonrewarding outcome will create a trace

of relatively small magnitude. Conversely, if the expectancy of reward is high, then a single rewarding outcome will create a relatively small trace and a single nonrewarding outcome will create a relatively large trace.

Decay of traces within the memory system. Probabilities of reward cannot be calculated directly from the initial trace magnitudes because earlier traces have decayed more than recent ones. For the purposes of this paper no theory of decay is necessary; however, it is tentatively assumed that simple decay with time adds to the effects of inteference from new learning to produce the phenomenon. Where both time intervals between trials and opportunities for new learning remain constant in a learning situation, it is proposed that each trace within the memory system decays by a constant fraction upon the occurrence of each new trace. Suppose the constant fraction were *arbitrarily* chosen as 10 per cent. A single memory trace of initial magnitude 100 might then decay, successively, from 100 to 90, to 81, to 73, to 66, to 59, etc. These new values represent the ***residual magnitudes*** remaining after successive reductions of initial trace magnitude through decay in the memory system. It follows from consideration of the decay function that recent events are weighted more heavily than less recent ones in the determination of probabilities upon which responses are based.

Selection of response. The organism weighs the "probabilities" of reward associated with each alternative and selects the alternative associated with the greatest "probability" of reward. As stated above, it seems unlikely that the calculus of a frequency theory of probability is used. It has been proposed that decay of traces within the memory system causes recent events to have greater weight than less recent ones in determining the probability estimates upon which responses are based. An approximation of the "subjective" probabilities which determine the organism's choice of response can be derived from the ***residual trace magnitudes*** discussed above. Three steps are involved. (*a*) An initial magnitude must be assigned to each behavior outcome. (*b*) A decay coefficient must be selected so that each memory trace may be reduced in magnitude upon the occurrence of each new trace. This determines residual trace magnitudes. (*c*) Positive ***residual*** magnitudes associated with a response are then summed to yield ΣPr. Negative ***residual*** magnitudes are summed to yield ΣNr. The probability of positive outcome associated with response (r) is then calculated:

$$p = \frac{\Sigma Pr}{\Sigma Pr + \Sigma Nr}$$

This model may be used for individual response prediction by calculating the probabilities associated with each alternative and then predicting that S will select the alternative associated with the greatest probability of reward. Predictions based upon this model will thus be determined jointly by frequency and recency of reward.

An illustration will make clear the differences between predictions based upon a recency-probability model and predictions based upon frequency or traditional probability calculations. Suppose, to take an extreme example, that S is rewarded on five successive trials and that he is not rewarded for the next five successive trials. For a second example, suppose that the first five trials are nonrewarded, but that a second five trials are rewarded. Frequency of reward, frequency of response, or an algebraic summation of positive and negative outcomes fails to distinguish between the first and second examples. Likewise, probability based upon relative frequency of reward is the same for the

two examples. The recency-weighted probability model, which has been tentatively proposed here, leads to very different "probabilities" in the two examples. To greatly simplify computation, suppose that the initial trace magnitude of each of the 10 outcomes was 100, and that a decay coefficient of 10 per cent is accepted. In the first example, where reward followed only the first five trials, the sum of positive residual traces is $\Sigma Pr = 242$, and the sum of negative residual traces is $Nr = 410$. The recency-weighted probability

$$p = \frac{\Sigma Pr}{\Sigma Pr + \Sigma Nr} = 0.37$$

In the second example only the last five of the 10 responses resulted in reward. In this case $\Sigma Pr = 410$ and $\Sigma Nr = 242$. The recency-weighted probability in the second example is

$$p = \frac{\Sigma Pr}{\Sigma Pr + \Sigma Nr} = 0.63$$

The recurrence of the response is more likely where the last five of ten responses were rewarded than where only the first five of ten were rewarded.

It should be pointed out that the stochastic models of Estes and Burke (5) and Bush and Mosteller (1) incorporate several of the principles emphasized in this paper. Estes and Burke quantify association theory in such a manner that predictions derived from their model are based upon both recency and frequency. Also, in common with the cognitive formulation they state, intuitively, that a reinforced trial following a sequence of nonreinforced trials adds a larger associative increment than one reinforced trial in a sequence of reinforced trials. This function has been assigned to expectancy-based differences in initial trace magnitudes in the cognitive probability model. The Bush and Mosteller model is similar in many respects to that of Estes and Burke; it differs from the latter model primarily in the functions assigned to reinforcement and inhibition. It appears likely that both models could lead to accurate response predictions in the experimental situations discussed below. Neither of these theories was utilized as the basis for individual response predictions in the following experiment because each requires the use of an arbitrary (until empirically established) weighting coefficient. If any invidious comparisons were to result from predictions the present writers might derive from these models, it is quite probable that selection of other coefficients would substantially change the results. Further extension of these models by their authors into the area of individual response predictions is eagerly awaited.

CHOICE OF EXPERIMENTAL SITUATIONS

In a great many situations employed in typical learning experiments, it is impossible to assess the importance of *order* of events in the learning sequence. It may be for this reason that recency has been neglected in theoretical treatments of learning. To evaluate the effects of recency, a situation is required in which one response will occur if frequency is the determining factor and another response will occur if recency is the important variable. A selective learning situation such as a simple T maze should provide the necessary setting. Nevertheless, it should be noted that in many T-maze situations frequency and recency lead most often to identical predictions. Where reward is placed in one goal box on a high proportion of the trials, the animal soon responds predominantly to that side; consequently, the more frequent and most recent response are the same. A T-maze situation in which reward is randomly varied between the two alternatives at a 50–50 ratio should provide a

maximum of response variation. In such a situation the difference between frequency and recency should be most evident, and the predictive efficacy of the various theoretical models should be put to the most difficult test.

EXPERIMENT I

This experiment was designed to compare the relative predictive efficacy of the theoretical models discussed above.

Procedure

Fifteen albino rats of the Sprague-Dawley strain were given 25 trials per day for two consecutive days on an elevated T maze. The noncorrection method was used. All animals were experimentally naive at the beginning of the experiment, and each was given pretraining on a straight runway. During the experimental trials, reward was randomly varied between the alternative goal boxes in a 50–50 ratio. The schedule of reward placement was constructed from a table of random numbers.

The maze was 4 ft. long by 2 in. wide with a starting arm 2 ft. by 2 in. The maze and two similar goal boxes were painted a uniform gray. In order to reduce distracting extra-maze cues, the experiment was conducted in the "dome" room of the Comparative Psychology Laboratory at the University of Texas. The experimental dome is 18 ft. in diameter with 5.5-ft. side walls arched to an 8-ft. center. It is constructed symmetrically of wooden ribs and covered with a fine mesh screen painted silver gray, hence giving a homogeneous surface radially symmetrical about the geometric center.

Analysis of Results

The problem was designed to determine which of the theoretical models discussed above most accurately predicts the individual responses of the fifteen *S*s in this experiment. Since frequency and probability models require a minimal amount of past experience upon which prediction can be based, no predictions were made for the first five responses of each day. Thus, predictions were calculated for a total of 40 responses for each of the fifteen *S*s. Table 1 compares the accuracy of predictions based upon each theoretical model.

The values entered in Column "*t*" of this table are *t*-test values for the differences between the mean accuracy of predictions based upon the specified model and the mean accuracy of predictions derived from the recency-weighted model. Since it was predicted that the latter theoretical model would prove to be a more accurate predictor of individual responses than any of the other models considered, a one-tailed test is appropriate. (See Jones, 8.) A t of 1.701 is significant at the .05 level, and a t of 2.467 is significant at the .01 level of confidence where 28 degrees of freedom are available. Predictions based upon the recency-weighted probability model were made by employ-

TABLE 1

AVERAGE ACCURACY OF RESPONSE PREDICTIONS

Source of Prediction	Mean Accuracy	*SD*	*t*
1. Postremity—"go where went last trial"	65.1	12.6	2.46
2. Postremity—"go where went last if rewarded, but alternate if not rewarded"	67.2	7.3	2.55
3. Frequency of response	62.2	15.5	2.75
4. Frequency of reward	64.8	13.8	2.06
5. Reaction potential—algebraic sum of positive and negative outcomes	61.2	10.7	3.75
6. Probability $\left(p = \dfrac{\Sigma P}{\Sigma P + \Sigma N}\right)$	67.2	11.3	2.11
7. Recency-weighted probability model	75.2	9.7	—

ing two gross approximations. The only excuses for this procedure are found in the accuracy of predictions actually achieved, and in the likelihood that more precise coefficients should lead to even more accurate predictions. The first approximation involved the assignment of initial trace magnitudes to events in the learning sequence. The first rewarded or nonrewarded outcome in a succession of consecutive rewarded or nonrewarded outcomes was arbitrarily assigned a value of 100. Each successive reward following reward, or nonreward following nonreward, was arbitrarily assigned an initial trace magnitude which was 10 per cent less than the magnitude of the trace preceding it. When there was a shift from reward to nonreward or from nonreward to reward, the first outcome in the new sequence was assigned an initial magnitude of 100, and each successive outcome in this new run of reward or nonreward was reduced by 10 per cent. The second arbitrary approximation involved the selection of a 10 per cent decay coefficient. Upon the occurrence of each new outcome the magnitude of each residual trace was assumed to decay to 90 per cent of what it was before the occurrence of the new outcome. Through the accumulation of empirical evidence, it should be possible to correct and refine coefficients and to secure significantly more accurate predictions than those which resulted from this approximate method; nevertheless, a significant increase in accuracy has resulted from employing even this rough approximation of the theoretical model.

EXPERIMENT II

A second experiment was designed to compare the relative efficacy of predictions based upon frequency and recency. Since no attempt was made to distinguish between postremity and the recency-weighted probability model tentatively proposed here, the results of this experiment may also be interpreted as a test of the postremity principle in individual response prediction.

Procedure

Thirty hooded rats were *Ss* in this experiment. Each was experimentally naive at the beginning of the experiment, and each was given pretraining on a straight runway. The apparatus was an elevated T maze 4 ft. in length with starting arm 2 ft. long. All units were 2 in. wide, and the maze surface was 15 in. from the floor. To eliminate extramaze cues, the experiment was conducted in the dome room described in Experiment I.

Each animal received all experimental trials on one day. Between each trial, *Ss* were returned to a restraining compartment for approximately 1 min. of rest. The noncorrection method was used throughout the experiment. On the experimental day, one trial was given to determine any initial preference; no reward was present in the maze during this trial. On succeeding training trials, reward was always present in one of the two goal boxes. A schedule of reward placement was constructed from a table of random numbers in such a manner that reward was predominantly on one side of the maze during the first half of training, but was then reversed so that it was presented predominantly on the other alternative for the latter half of training. Specifically, reward was presented in 9 out of the first 13 trials on the side initially preferred by the *S*, and then the schedule was reversed so that reward was presented in 9 out of the last 13 training trials on the side not initially preferred. Following the 26 trials in which reward was always present in one goal box, 15 additional trials were administered during which no reward was presented.

RESULTS AND CONCLUSIONS

Criterion Measure

An attempt was made to create a situation in which frequency and recency should lead to different predictions. Because reinforcement is postulated to have important effects upon the strengths of competing response tendencies, responses during extinction were selected as best for evaluating the

relative magnitudes of competing response tendencies.

The relative number of responses to each alternative during a complete "alternation cycle" has been utilized by Hull (7) as a measure of the relative strengths of computing reaction potentials. He states:

> . . . The momentary oscillation principle does not prevent the occurrence of appreciable sequences of one reaction to the exclusion of the other. Perhaps the most fundamental concept . . . is that of *response alternation*. A response alternation is said to occur when one type of response shifts to the other. For example, in the response-sequence fragment
>
> $$R- \,|R+ \; R+ \; R+|\, \mathrm{R}- \; R- \,|R+$$
>
> there are three alternations, each marked by vertical lines. Our second concept, flowing directly from the first, is that of the *alternation phase;* this includes the number of reactions falling between two successive response alternations. Thus, in the above example, the first complete alternation phase represented contains three R+'s. Our third concept is that of the *alternation cycle*. An alternation cycle is the succession of responses comprised in two successive alternation phases. In the above example, an alternation cycle of $3+2$, or five reactions, is enclosed between the two heavy vertical lines. Finally, there is the concept of *asymmetry* of the response cycle; this refers to the fact that a behavior cycle may contain more reactions in one alternation phase than in the other (7, p. 42).

It is the asymmetry of responses during the first complete alternation cycle of the extinction period that is of especial interest to us here. Hull shows that where the reaction potential associated with one response is greater than that associated with the other, response alternation cycles will be asymmetrical, with a majority of responses made to the alternative associated with the greater reaction potential. Thus, the relative length of phases during an alternation cycle is an acceptable criterion for inferring relative strengths of competing response tendencies. Predictions based upon frequency and recency principles were made for responses of the first complete alternation cycle of the extinction period, and a chi-square test used to determine the significance of number of correct predictions.

Results

Twenty-one of the 30 *S*s made a majority of responses, during the first complete alternation cycle of the extinction period, to the alternative on which *reward* had been most *recently* encountered. A chi square of 4.033 with 1 df is significant at the .05 level of confidence.

If it is predicted that a majority of responses during the first complete alternation cycle will be made to the alternative *responded* to most recently during training, it is found that the behavior of 22 of the 30 *S*s is correctly predicted. With a one-tailed test, which is appropriate in light of the theoretical hypothesis of this paper, a chi square of 5.633 with 1 df is significant at the .01 level of confidence.

When it is predicted that a majority of responses during the first complete alternation cycle will be to the side *responded* to most *frequently,* the behavior of only 16 out of 30 *S*s is correctly predicted.

The behavior of 17 out of 30 *S*s is

TABLE 2

RECENCY AND FREQUENCY PRINCIPLES IN PREDICTING MAJORITY OF RESPONSES DURING FIRST ALTERNATION CYCLE OF EXTINCTION PERIOD

Predict Majority of Responses	Correct	Incorrect
1. To alternative responded to most *recently*	22	8
2. To alternative rewarded most *recently*	21	9
3. To alternative responded to most *frequently*	16	14
4. To alternative rewarded most *frequently*	17	13

correctly predicted on the basis of *frequency of reward*. Because the total number of reward placements was equated for the two alternatives, the algebraic sum of positive and negative outcomes associated with responses to each side was always equal. Such a model would lead to the prediction of an equal probability for each alternative or, in other words, to no directional prediction in this situation.

SUMMARY

1. One of the important controversies among contemporary learning theorists is the question of whether *S*s utilize all past experiences equally or whether recent experiences are of greater importance in determining responses in a selective learning situation.

2. It is possible to utilize contemporary theories of learning as a basis for individual response predictions. Because learning is a change within the individual organism, the final test of a theory of learning is its ability to predict individual behavior.

3. The work of Voeks indicates that postremity predictions are less accurate where they conflict with frequency-based predictions than where the two predictions are not opposed, although postremity predictions were more accurate on the average in her situation. This suggests that a theoretical model which bases predictions upon *both* frequency and recency should be superior to one which utilizes only one principle or the other.

4. After reviewing models for individual response predictions which may be derived from postremity, frequency, and probability theories, a predictive model was tentatively suggested which incorporated both recency and frequency into a cognitive theory of learning.

5. Experiment I presented a T-maze situation in which reward was randomly varied (50–50) between the two alternatives. Individual responses were predicted from postremity, frequency, and probability models. The accuracy of these predictions was compared to the accuracy of predictions based upon a model taking both frequency and recency into consideration. The latter model was found to predict with significantly higher accuracy than were any of the other models from which predictions were derived.

6. Expriment II presented a T-maze situation in which the reinforcement schedule first favored one alternative and was then reversed to favor the other alternative. The learning period was followed by an extinction period. The relative number of responses made to each alternative during the first complete alternation cycle of the extinction period was used as a measure of the relative strengths of competing response tendencies developed during the learning period. In this situation, it was found that recency predicted significantly better than chance, while frequency predicted only at the chance level.

7. The general conclusion of this paper is that a theoretical model which incorporates both the principles of recency and frequency can predict individual responses more accurately than one which considers only recency or frequency alone. Estes and Burke (5) have demonstrated that Guthrie's contiguous association theory can be quantified to serve as the basis for predictions which are based upon both recency and frequency. It should be possible for any of the contemporary theories of learning to incorporate both principles. Until frequency and probability theories include a recency principle it appears that they will be neglecting an important consideration, as shown by the growing body of data on the relative importance of recent events in the learning sequence.

8. This paper has indicated how a

cognitive-probability model might incorporate both recency and frequency principles. It is also suggested that refinement of predictions based upon this model can be achieved through developing coefficients which more accurately reflect empirical data.

REFERENCES

1. BUSH, R. R., & MOSTELLER, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, **58**, 313–323.
2. CRESPI, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.
3. CRESPI, L. P. Amount of reinforcement and level of performance. *Psychol. Rev.*, 1944, **51**, 341–357.
4. ELLIOTT, M. H. The effect of change of reward on the maze performance of rats. *Univer. Calif. Publ. Psychol.*, 1928, **4**, 19–30.
5. ESTES, W. K., & BURKE, C. J. Stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
6. GUTHRIE, E. R. *The psychology of learning.* New York: Harper, 1935.
7. HULL, C. L. *A behavior system: an introduction to behavior theory concerning the individual organism.* New Haven: Yale Univer. Press, 1952.
8. JONES, L. V. Test of hypotheses: one-sided vs. two-sided alternatives. *Psychol. Bull.*, 1952, **49**, 43–46.
9. SPENCE, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.
10. TINKLEPAUGH, O. L. An experimental study of representative factors in monkeys. *J. comp. Psychol.*, 1928, **8**, 197–236.
11. TOLMAN, E. C. *Purposive behavior in animals and man.* New York: The Century Co., 1932.
12. TOLMAN, E. C., & BRUNSWIK, E. The organism and the causal texture of the environment. *Psychol. Rev.*, 1935, **42**, 43–77.
13. VOEKS, V. W. Postremity, recency, and frequency as basis for prediction in the maze situation. *J. exp. Psychol.*, 1948, **38**, 495–510.
14. VOEKS, V. W. A formalization and clarification of a theory of learning. *J. Psychol.*, 1950, **30**, 341–362.

(Received March 30, 1957)

Psychological Review
Vol. 64, No. 5, 1957

AN ANALYSIS OF STIMULUS VARIABLES INFLUENCING THE PROPRIOCEPTIVE CONTROL OF MOVEMENTS

HARRY P. BAHRICK

Ohio Wesleyan University [1, 2]

It is generally known that accurate execution of movements depends upon proprioceptive information reaching the central nervous system. Clinical evidence (**7**, p. 235) as well as experimental findings (**8**) indicates that control and perception of movements are very poor when this sensory channel is not functioning.

Little is known, however, about the specific characteristics of proprioceptive stimulation that permit the individual to control changes in position, rate, or acceleration of his limbs. In other words, no detailed theories of proprioception comparable to the specific theories available for some other sense modalities have been developed (**7**, p. 234). Most of the available knowledge in this area is based upon anatomical investigations of the receptor system, its neural connections, and its central representations. Although several types of receptors have been identified (**13**, p. 1185; **14**), differentiation of their function is as yet not clearly established. It is thought that forces internal to the body act as proprioceptive stimuli, but the processes by which these stimuli are encoded into messages which ultimately form the basis for perception and control of movements are not well understood (**7**, p. 234).

Behavioral data specifying the relations between stimulus and response characteristics have been difficult to obtain because of problems of controlling proprioceptive stimuli. Investigators have used drugs or faradic currents (**8**, **9**) as means of reducing the effectiveness of proprioceptive stimuli. Recently, an indirect approach to this problem has been attempted. This approach consists of varying the type and degree of resistance to motion offered by a control which *S* uses in the execution of movements. The effect of this variation upon *S*'s ability to perceive and control his movements is studied, and an attempt is made to infer characteristics of the proprioceptive system. As a technique of investigating proprioception, this approach has obvious limitations. The forces which *S* applies to move a control are only indirectly related to the proprioceptive stimulation he receives during the execution of the movement. The cutaneous senses are also stimulated during movement, and unknown transformations are involved between the control force acting upon the limb and the proximal stimuli acting upon receptors in muscles, tendons, or joints.

Despite these substantial limitations, the approach has some theoretical as well as practical advantages. The forces required to move a control, and thus also the control forces acting upon the

[1] This research was supported in part by the United States Air Force under Contract No. AF 41(657)-70 monitored by the Operator Laboratory, Air Force Personnel and Training Research Center, Randolph Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[2] Work contributing to this theory was initiated in the Laboratory of Aviation Psychology of the Ohio State University in 1953, and is now being continued there. For the past year the author has continued research contributing to this theory at Ohio Wesleyan University through support from the National Science Foundation.

limb, can be specified as a function of four physical properties of the control, and these properties can all be regulated conveniently by *E*. They are mass, viscosity, elasticity, and the degree of coulomb friction. In a control such as that used by Howland and Noble (11) these parameters combine according to the following time-varying system equation:

$$L_t = K\theta + B\,d\theta/dt + J\,d^2\theta/dt^2, \quad (1)$$

where the left-hand side of the equation is the force applied by the human arm, the right-hand side represents the component resistive forces offered by the external control, L_t is the torque required to move the control at any instant of time (t), K is the constant of elasticity of the control, B is the viscosity constant, J is the moment of inertia, and θ is the angular displacement of the control with respect to its neutral, or spring-centered, position (6). Coulomb friction has been neglected in this equation, as have the internal resistive forces in the limb itself.

It has already been pointed out that the force which the control exerts upon the limb may not be equated to proprioceptive stimuli. However, one may assume that the forces which do act as proprioceptive stimuli during the movement of limbs are determined by physical properties of our limbs analogous to those specified in the above equation for the control. Previous investigation (4) has already established some of these physical properties of limbs and their significance in relation to the control of movements. These physical properties of limbs are difficult to control, and the present approach attempts to infer their function in proprioception by studying the effects of analogous characteristics of controls where these properties can be manipulated conveniently.

From an applied viewpoint, this approach may be useful in that the data are relevant to the solution of human engineering problems related to the design of controls used in man-machine systems, or to the design of prosthetic devices.

In the present article some general hypotheses are developed about the effect of each of the physical control parameters specified in Equation 1, and data are reported which test these hypotheses.

Inspection of Equation 1 shows that the torque needed to move the control depends upon its position, rate, and angular acceleration, the relative importance of these depending upon the respective values of the elasticity, damping, and inertia constants. Thus, if the elasticity constant K is zero, the torque required to move the control will be independent of its position, but if K is relatively large, the torque will vary largely as a function of position. Analogous relations exist between the damping constant B and angular velocity, and between the moment of inertia and angular acceleration.

It is now hypothesized that a man can use the force cues obtained in moving the control to improve his perception of position, rate, and acceleration of limb motion. Specifically, it is hypothesized that the elasticity constant of the control improves *S*'s ability to perceive and control positions, the damping constant improves perception and control of rate, and the moment of inertia improves the perception and control of acceleration. Thus, an increase in each of these control constants should lead to improvement in the corresponding behavior. At the same time, it is hypothesized that an increase in any of the control constants will affect adversely performance which is aided by the other constants. Thus, increases in K are expected to interfere with the control and perception of rate and acceleration, while increases in B and J will affect adversely the control and perception of position. This hypothesis suggests itself, since the force

required to move the control would not be expected to provide useful cues for the control of rate if it changes rapidly with position, and, conversely, it should not offer useful cues for the detection of position if it varied greatly as a function of rate or acceleration.

Several experiments have been conducted in which the accuracy of movement was studied as a function of the physical characteristics of controls (1, 2, 3, 10, 11). A few of these (1, 2, 10) were designed specifically to test the above predictions. In one study (2) *Ss* performed simple circular and triangular control motions with a joystick control which was loaded with various degrees of spring stiffness, or damping, or mass. In each control-loading condition, the movements were first practiced with the help of a visual guide and paced by means of a metronome. The visual and auditory guides were then removed, and *Ss* were instructed to reproduce the motions as accurately as possible. Photographic records of all motions were obtained and measured for accuracy of temporal and spatial reproduction. It was found that an increase of viscous damping or of inertia of the control resulted in greater uniformity of speed within individual motions, and also in greater uniformity of speed in successive reproductions of the same motion. In the case of the triangular motions, increased mass and increased damping led to greater uniformity of peak velocity on each side of the triangle on successive trials. Spring loading interfered with the control of rate and acceleration, but its effect upon spatial accuracy of the reproduced motion was, in general, not significant. It was suggested that extended practice is needed for effective utilization of cues provided by spring loading.

This hypothesis was checked in a second experiment (1) in which the accuracy of positioning a horizontal arm control was investigated as a function of changes in the torque-displacement relation of the control. Extended practice was given and knowledge of results was provided. It was found that positioning errors are smallest when the ratio of relative torque change to displacement is largest. Under optimum conditions of spring loading, average positioning errors were less than half the amount obtained for a control which was not spring loaded. It was concluded that force cues provided by a spring-loaded control can improve the accuracy of positioning a control, and that the amount of improvement is a function of the relative and absolute torque change per unit of amplitude change.

Further investigation [3] of the usefulness of force cues in regulating the amplitude of motion has supported the above conclusions. It was shown that the transmission of amplitude information can be increased significantly by spring loading the control used by *Ss*. Optimum results were obtained with a control which provided geometric increments of force as a function of arithmetic changes of amplitude. This condition provides force cues which are equally discriminable over the range of amplitudes employed (12, 15), and yields the largest number of absolutely discriminable categories of amplitude response.

Although the above results support the general hypotheses regarding the effects of *K, B,* and *J* constants upon the control of movements, many questions remain unanswered. In order to establish that the observed effects are due to changes in proprioceptive stimulation, it will be necessary to control cutaneous sensitivity. It is hypothesized here that

[3] Bahrick, H. P. Force cues and the control of movement amplitudes. (In preparation.)

the contribution of cutaneous receptors is most significant in relation to minute manipulatory responses, and least significant for larger movements of the type dealt with here.

Further problems arise because the control parameters under discussion have certain mechanical effects upon the nature of movements, and these must be separated from the effects upon proprioceptive stimulation. Large amounts of damping, for example, make rapid movements difficult and fatiguing, and greater uniformity of movement rate observed under these conditions may reflect mechanical effects rather than improved proprioceptive discrimination. The identification of these mechanical effects becomes more difficult when continuous movements are dealt with, as was shown in the study by Howland and Noble (**11**). Interactions among the physical parameters of controls may cause complex mechanical effects such as oscillation, and these may obscure or counteract the effects due to augmented proprioceptive stimulation. In general, the analysis of *K, B,* and *J* effects is relatively simple for discrete, adjustive movements of the type primarily dealt with so far, but becomes increasingly complicated for complex or continuous motions.

Work now in progress attempts to establish relations between the forces exerted upon a control, and intensity of stimulation at the receptors in the elbow. This analysis is based upon a simplified mechanical model of the arm (**16**) by means of which forces acting upon the hand are resolved at the elbow joint (**5**, p. 319). In this manner it may become possible to infer changing intensities of stimulation of receptors at the joint during the course of movements.

Ultimately, the development of proprioceptive theory described here must be supported by a more direct analysis of *K, B,* and *J* factors within the body, and their effects upon the perception and control of movements. This, in turn, will require a better understanding of the biophysical principles by which forces internal to the body are brought to bear upon proprioceptive receptors.

REFERENCES

1. BAHRICK, H. P., BENNETT, W. F., & FITTS, P. M. Accuracy of positioning responses as a function of spring loading in a control. *J. exp. Psychol.,* 1955, **49**, 437–444.
2. BAHRICK, H. P., FITTS, P. M., & SCHNEIDER, R. The reproduction of simple movements as a function of proprioceptive feedback. *J. exp. Psychol.,* 1955, **49**, 445–454.
3. DERWORT, A. Ueber die Formen unserer Bewegungen gegen verschiedenartige Widerstaende und ihre Bedeutung fuer die Wahrnehmung von Kraefton. *Z. f. Sinnesphysiol.,* 1943, **70**, 135–183.
4. FENN, W. O. The mechanics of muscular contraction in man. *J. appl. Physics,* 1938, **9**, 165–177.
5. FICK, R. *Handbuch der Anatomie und Mechanik der Gelenke.* Part II. Jena: Verlag von Gustav Fischer, 1910.
6. FITTS, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 1287–1340.
7. GELDARD, F. A. *The human senses.* New York: Wiley, 1953.
8. GOLDSCHEIDER, A. Untersuchungen ueber den Muskelsinn. I. Ueber die Bewegungsempfindung. In A. Goldscheider, *Gesammelte Abhandlungen,* Vol. II. Leipzig: Barth, 1898.
9. GOLDSCHEIDER, A. Untersuchungen ueber den Muskelsinn. II. Ueber die Empfindung der Schwere und des Widerstandes. In A. Goldscheider, *Gesammelte Abhandlungen,* Vol. II. Leipzig: Barth, 1898.
10. HELSON, H., & HOWE, W. H. Inertia, friction, and diameter in handwheel tracking. OSRD Rep. No. 3454, 1943. (PB 406114.)
11. HOWLAND, D., & NOBLE, M. E. The effect of physical constants of a control

on tracking performance. *J. exp. Psychol.*, 1953, **46**, 353–360.

12. JENKINS, W. L. The discrimination and reproduction of motor adjustment with various types of aircraft controls. *Amer. J. Psychol.*, 1947, **60**, 397–406.
13. JENKINS, W. L. Somesthesis. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1172–1190.
14. MATHEWS, B. H. C. Nerve endings in mammalian muscle. *J. Physiol.*, 1933, **78**, 1–53.
15. NOBLE, M. E., & BAHRICK, H. P. Response generalization as a function of intratask response similarity. *J. exp. Psychol.*, 1956, **51**, 405–412.
16. WHITE, H. E. *Modern college physics*. New York: Van Nostrand, 1948.

(Received for early publication June 6, 1957)

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by James Drever

Vol. 48 Part 3 August 1957 20s. net

The subscription price per volume, payable in advance, is 60s. (or $10) net (post free).

Subscriptions may be sent to any bookseller or to the

CAMBRIDGE UNIVERSITY PRESS
Bentley House, Euston Road, London, N. W. 1